

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Hisayoshi NAGAE, et al.

GAU:

SERIAL NO: New Application

EXAMINER:

FILED: Herewith

FOR: VARIABLE VOICE RATE APPARATUS AND VARIABLE VOICE RATE METHOD

REQUEST FOR PRIORITY

COMMISSIONER FOR PATENTS  
ALEXANDRIA, VIRGINIA 22313

SIR:

☐ Full benefit of the filing date of U.S. Application Serial Number , filed , is claimed pursuant to the provisions of 35 U.S.C. §120.

☐ Full benefit of the filing date(s) of U.S. Provisional Application(s) is claimed pursuant to the provisions of 35 U.S.C. §119(e): Application No. Date Filed

☒ Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of 35 U.S.C. §119, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

COUNTRY

Japan

APPLICATION NUMBER

2002-382385

MONTH/DAY/YEAR

December 27, 2002

Certified copies of the corresponding Convention Application(s)

☒ are submitted herewith

☐ will be submitted prior to payment of the Final Fee

☐ were filed in prior application Serial No. filed

☐ were submitted to the International Bureau in PCT Application Number

Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.

☐ (A) Application Serial No.(s) were filed in prior application Serial No. filed ; and

☐ (B) Application Serial No.(s)

☐ are submitted herewith

☐ will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,  
MAIER & NEUSTADT, P.C.



Marvin J. Spivak

Registration No. 24,913

Customer Number

22850

Tel. (703) 413-3000  
Fax. (703) 413-2220  
(OSMMN 05/03)

C. Irvin McClelland  
Registration Number 21,124

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日            2 0 0 2 年 1 2 月 2 7 日  
Date of Application:

出 願 番 号            特 願 2 0 0 2 - 3 8 2 3 8 5  
Application Number:  
[ST. 10/C]:            [ J P 2 0 0 2 - 3 8 2 3 8 5 ]

出      願      人            株 式 会 社 東 芝  
Applicant(s):

2 0 0 3 年   7 月   8 日

特許庁長官  
Commissioner,  
Japan Patent Office

太田信一郎



【書類名】 特許願

【整理番号】 A000204778

【提出日】 平成14年12月27日

【あて先】 特許庁長官 殿

【国際特許分類】 G10L 3/00

【発明の名称】 話速可変装置及び話速変換方法

【請求項の数】 18

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

【氏名】 永江 尚義

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

【氏名】 桃崎 浩平

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

【氏名】 蟻生 政秀

【特許出願人】

【識別番号】 000003078

【氏名又は名称】 株式会社 東芝

【代理人】

【識別番号】 100058479

【弁理士】

【氏名又は名称】 鈴江 武彦

【電話番号】 03-3502-3181

## 【選任した代理人】

【識別番号】 100084618

【弁理士】

【氏名又は名称】 村松 貞男

## 【選任した代理人】

【識別番号】 100068814

【弁理士】

【氏名又は名称】 坪井 淳

## 【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

## 【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

## 【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

## 【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

## 【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

## 【提出物件の目録】

【物件名】 明細書 1

【物件名】

図面 1

【物件名】

要約書 1

【プルーフの要否】

要

【書類名】 明細書

【発明の名称】 話速可変装置及び話速変換方法

【特許請求の範囲】

【請求項 1】 音声の再生速度を可変制御する話速可変装置において、  
音声信号から音声データを生成する音声データ生成手段と、  
この音声データ生成手段が生成した音声データに基づいて、この音声データが  
有する発声内容を示すテキストデータを生成するためのテキストデータ生成手段  
と、  
このテキストデータを、言語的に意味を持つ単位に分割する分割情報を生成す  
る分割情報生成手段と、  
音声の再生制御に関わる前記単位に設定される再生情報を生成する再生情報生  
成手段と、  
前記再生情報と前記分割情報とに基づいて、前記音声データから前記単位で音  
声再生を制御する音声再生制御手段とを具備することを特徴とする話速可変装置  
。

【請求項 2】 音声の再生速度を可変制御する話速可変装置において、  
音声信号から音声データを生成する音声データ生成手段と、  
この音声データ生成手段が生成した音声データに基づいて、この音声データが  
有する発声内容を示すテキストデータを生成するためのテキストデータ生成手段  
と、  
このテキストデータを、言語的に意味を持つ単位に分割する分割情報を生成す  
る分割情報生成手段と、  
音声の再生制御に関わる前記単位に設定される再生情報を生成する再生情報生  
成手段と、  
前記再生情報を記憶する第 1 の記憶手段と、  
前記第 1 の記憶手段が記憶する再生情報と、前記分割情報とに基づいて、前記  
音声データから前記単位で音声再生を制御する音声再生制御手段とを具備するこ  
とを特徴とする話速可変装置。

【請求項 3】 音声の再生速度を可変制御する話速可変装置において、

音声信号から音声データを生成する音声データ生成手段と、

この音声データ生成手段が生成した音声データに基づいて、この音声データが有する発声内容を示すテキストデータを生成するためのテキストデータ生成手段と、

このテキストデータを、言語的に意味を持つ単位に分割する分割情報を生成する分割情報生成手段と、

音声の再生制御に関わる前記単位に設定される再生情報を生成する再生情報生成手段と、

前記再生情報を記憶する第1の記憶手段と、

前記音声データと分割情報とを記録する第2の記憶手段と、

前記第1の記憶手段が記憶する再生情報と、前記第2の記憶手段が記憶する分割情報とに基づいて、前記第2の記憶手段が記憶する音声データから前記単位で音声信号を生成する音声信号生成手段とを具備することを特徴とする話速可変装置。

【請求項4】 前記テキストデータ生成手段は、前記音声データ生成手段が生成した音声データを解析し、この音声データが有する発声内容を示すテキストデータにて記述することを特徴とする請求項1乃至請求項3のいずれかに記載の話速可変装置。

【請求項5】 前記テキストデータ生成手段は、

前記音声データ生成手段が生成した音声データに基づく音声をオペレータに対して拡声する拡声手段と、

前記オペレータの操作により、テキストデータの入力を受け付けるテキスト入力手段とを備えることを特徴とする請求項1乃至請求項3のいずれかに記載の話速可変装置。

【請求項6】 さらに、情報を入力する入力手段を備え、

前記再生情報生成手段は、前記入力手段から入力された情報を再生情報として生成することを特徴とする請求項1乃至請求項5のいずれかに記載の話速可変装置。

【請求項7】 前記再生情報生成手段は、前記再生情報として、前記単位の

種別に応じた情報を生成するものであって、

前記音声再生制御手段は、前記再生情報と前記分割情報とに基づいて、前記単位毎の再生時間を決定し、この決定した時間で各単位を再生する音声再生を前記単位で制御することを特徴とする請求項 1 乃至請求項 6 のいずれかに記載の話速可変装置。

【請求項 8】 前記分割情報生成手段は、前記テキストデータ生成手段により生成されたテキストデータを、品詞レベルで分割する分割情報を生成し、

前記再生情報生成手段は、再生情報として、前記単位の品詞に応じた重み付け情報を生成するものであって、

前記音声再生制御手段は、前記再生情報と前記分割情報とに基づいて、各単位に対して該当する品詞に応じた重み付け情報に基づく再生時間を決定し、この決定した時間で各単位を再生する音声再生を前記単位で制御することを特徴とする請求項 7 に記載の話速可変装置。

【請求項 9】 前記音声再生制御手段は、前記再生情報と前記分割情報とに基づいて、分割された単位のうち再生すべき単位を選択し、この選択された単位を再生する音声再生を前記単位で制御することを特徴とする請求項 1 乃至請求項 6 のいずれかに記載の話速可変装置。

【請求項 10】 前記再生情報生成手段は、前記再生情報として、値 M を設定するものであって、

前記音声再生制御は、前記再生情報と前記分割情報とに基づいて、分割された単位のうち、M 単位のうち 1 個の単位を選択し、この選択された単位を再生する音声再生を前記単位で制御することを特徴とする請求項 9 に記載の話速可変装置。

【請求項 11】 前記再生情報生成手段は、前記再生情報として、文字列を示す情報を生成するものであって、

前記音声再生制御は、前記再生情報と前記分割情報とに基づいて、分割された単位のうち、前記再生情報が示す文字列を含む単位を選択し、この選択された単位を再生する音声再生を前記単位で制御することを特徴とする請求項 1 乃至請求項 6 のいずれかに記載の話速可変装置。



【請求項 1 2】 前記再生情報生成手段は、前記再生情報として、文字列を示す情報を生成し、値Nとを設定するものであって、

前記音声再生制御手段は、前記再生情報と前記分割情報とに基づいて、分割された単位のうち、前記再生情報が示す文字列をのべN個以上含む単位を選択し、この選択された単位を再生する音声再生を前記単位で制御することを特徴とする請求項 1 1 に記載の話速可変装置。

【請求項 1 3】 前記再生情報生成手段は、前記再生情報として、文字列を示す情報と、これに対応づけた優先度を示す情報とを生成するものであって、

前記音声再生制御手段は、前記再生情報と前記分割情報とに基づいて、分割された単位のうち、前記再生情報が示す優先度の高い文字列を含む単位を優先的に選択し、この選択された単位を再生する音声再生を前記単位で制御することを特徴とする請求項 1 乃至請求項 6 のいずれかに記載の話速可変装置。

【請求項 1 4】 さらに、各単位を再生するのに要する時間をそれぞれ検出する検出手段を備え、

前記音声再生制御手段は、前記再生情報と前記分割情報と、前記検出手段が検出した時間とに基づいて、分割された単位のうち、前記再生情報が示す優先度の高い文字列を含み、かつ再生に要する時間の合計が所定時間以内になる単位を優先的に選択し、この選択された単位を再生する音声再生を前記単位で制御することを特徴とする請求項 1 3 に記載の話速可変装置。

【請求項 1 5】 前記再生情報生成手段は、前記再生情報として、所定の単位が所定の順序で組み合わさる確率を示す情報を記憶するものであって、

前記音声再生制御手段は、前記再生情報と前記分割情報とに基づいて、分割された単位のうち、前記確率が予め設定した閾値より低い組み合わせの単位を選択し、この選択された単位を再生する音声再生を前記単位で制御することを特徴とする請求項 1 乃至請求項 6 のいずれかに記載の話速可変装置。

【請求項 1 6】 さらに、各単位を再生するのに要する時間をそれぞれ検出する検出手段を備え、

前記音声再生制御手段は、前記再生情報と前記分割情報とに基づいて、分割された単位のうち、前記確率が予め設定した閾値より低く、かつ再生に要する時間

の合計が所定時間以内になる組み合わせの単位を選択し、この選択された単位を再生する音声再生を前記単位で制御することを特徴とする請求項 15 に記載の話速可変装置。

【請求項 17】 音声の再生速度を可変制御する話速可変装置において、音声の再生制御に関わる前記単位に設定される再生情報を生成する再生情報生成手段と、

前記再生情報と、入力される音声信号の音声データから、言語的に意味を持つ単位に分割した分割情報とに基づいて、前記音声データから前記単位で音声再生を制御する音声再生制御手段とを具備することを特徴とする話速可変装置。

【請求項 18】 音声の再生速度を可変制御する話速可変方法において、音声信号から音声データを生成するステップと、この音声データ生成手段が生成した音声データに基づいて、この音声データが有する発声内容を示すテキストデータを生成するステップと、このテキストデータを、言語的に意味を持つ単位に分割する分割情報を生成するステップと、

音声の再生制御に関わる前記単位に設定される再生情報を生成するステップと、前記再生情報と前記分割情報とに基づいて、前記音声データから前記単位で音声再生を制御するステップとを具備することを特徴とする話速可変方法。

#### 【発明の詳細な説明】

##### 【0001】

##### 【発明の属する技術分野】

この発明は、例えば映像機器、音響機器、医療機器などの再生音声の速度を可変するために用いられる話速可変装置に関する。

##### 【0002】

##### 【従来の技術】

近年、音声、あるいは映像と関連付けられた音声を取り扱う機器の主流は、従来の磁気テープから光磁気ディスクや半導体メモリなどに移り、その記録方式もアナログ方式からデジタル方式に移行するようになった。

## 【0 0 0 3】

このように音声の記録方式がデジタル化されることで、再生速度の制御も容易になった。記録時よりも高速な再生を行う早聞き再生により同じ内容の音声を短時間で視聴したり、逆に遅く再生するスロー再生により丁寧に聞き取ることができる。

## 【0 0 0 4】

最近のオーディオ機器やパソコン上の音声再生アプリケーションソフトウェアの一部には、デジタル音声の特徴を利用した再生速度の可変機能が搭載されている。

## 【0 0 0 5】

この機能は、音声の波形データの特徴に基づいて制御する方式が使用されており、デジタル化された音声の波形データそのものについて一律に時間軸を短縮したり、分析したパラメータ上で全体の時間軸を短縮して再度合成したり、ある時間単位で飛び飛びに再生したりする方法が用いられている。さらには無音部分を検出してそれを再生対象から削除することで全体の時間を短縮する方法も用いられる（例えば、特許文献 1 参照）。

## 【0 0 0 6】

しかし、上述した従来の波形や分析パラメータ上で全体の時間軸を縮める方法では、縮めすぎると人間が聞き取れなくなるために、あまり高倍速にすることはできない。

## 【0 0 0 7】

また飛び飛びに再生する方法では、音声の人が喋る言葉の場合に意味が理解できなくなることがある。そして、無音部分を削除する方法では、音声中の無音部分が少ない場合に効果が得られない場合がある。

## 【0 0 0 8】

以上のように、従来の手法では、音声の音響的な特徴に基づいた制御を行っているため、再生される音声を示す内容を理解するのが困難な場合が多いという問題があった。

## 【0 0 0 9】

**【特許文献 1】**

特許第 3219892 号公報（第 4-9 頁、図 1）

**【0010】****【発明が解決しようとする課題】**

従来では、音声の音響的な特徴に基づいた制御を行っているため、再生される音声が表示内容を理解するのが困難な場合が多いという問題があった。

この発明は上記の問題を解決すべくなされたもので、再生速度を可変したり、再生時間を短縮しても、再生される音声が表示内容を理解しやすい話速可変装置及び話速変換方法を提供することを目的とする。

**【0011】****【課題を解決するための手段】**

上記の目的を達成するために、この発明は、音声の再生速度を可変制御する話速可変装置において、音声信号から音声データを生成する音声データ生成手段と、この音声データ生成手段が生成した音声データに基づいて、この音声データが有する発声内容を示すテキストデータを生成するためのテキストデータ生成手段と、このテキストデータを、言語的に意味を持つ単位に分割する分割情報を生成する分割情報生成手段と、音声の再生制御に関わる単位に設定される再生情報を生成する再生情報生成手段と、再生情報と分割情報とに基づいて、音声データから単位で音声再生を制御する音声再生制御手段とを具備して構成するようにした。

**【0012】**

またこの発明では、音声の再生速度を可変制御する話速可変方法において、音声信号から音声データを生成するステップと、この音声データ生成手段が生成した音声データに基づいて、この音声データが有する発声内容を示すテキストデータを生成するステップと、このテキストデータを、言語的に意味を持つ単位に分割する分割情報を生成するステップと、音声の再生制御に関わる単位に設定される再生情報を生成するステップと、再生情報と分割情報とに基づいて、音声データから単位で音声再生を制御するステップとを備えて構成するようにした。

**【0013】**

上記構成の話速可変装置及び話速変換方法では、音声信号から生成した音声データをテキストデータに変換し、このテキストデータを言語的に意味を持つ単位に分割する分割情報を生成し、音声の再生制御に関わる再生情報と上記分割情報とに基づいて、音声データから上記単位で音声信号を生成するようにしている。

#### 【0014】

したがって、上記構成の話速可変装置及び話速変換方法によれば、言語的に意味を持つ単位で音声信号を生成できるので、上記単位毎の制御により再生時間を短縮したり再生速度を高速化しても、再生される音声が表示内容を理解しやすい。

#### 【0015】

##### 【発明の実施の形態】

以下、図面を参照して、この発明の一実施形態について説明する。

図1は、この発明の第1の実施形態に係わる話速可変装置の構成を示すものである。

#### 【0016】

この図に示す話速可変装置は、マイクロホン100、音声データ生成部101、テキストデータ生成部102、単位分割部103a、入力部104、音声再生制御部105aおよびスピーカ106を備えている。

#### 【0017】

音声データ生成部101は、マイクロホン100から入力されるアナログ音声信号を音声データに変換し、これをテキストデータ生成部102および音声再生制御部105aに出力する。

#### 【0018】

テキストデータ生成部102は、音声認識技術を駆使して、音声データ生成部101から入力された音声データに基づいて、その発声内容を解析してテキストデータに変換する。

#### 【0019】

単位分割部103aは、テキストデータ生成部102で得られたテキストデータを、言語的に意味を持つ単位に分割し、各単位に再生情報を設定し、これをテ

キスト単位データとして音声再生制御部 105a に出力する。

#### 【0020】

ここで、言語的に意味を持つ単位とは、例えば「形態素」、「単語」、「文節」、「連文節」、「文」、「段落」のいずれかを意味する。そして、再生情報とは、分割された単位の種別（例えば品詞）に応じた重み付けの値  $\alpha$  である。

#### 【0021】

入力部 104 は、例えばオペレータや外部機器から、音声再生の速度に関わる情報として、再生倍率  $N$  と総再生時間  $L$  の入力を受け付けるものである。ここから入力された情報は、後述する音声再生制御部 105a 内のメモリに記憶される。

#### 【0022】

ここで、等倍速再生を行う場合には「 $N=1$ 」、一時停止する場合には「 $N=0$ 」、スロー再生を行う場合には「 $0 < N < 1$ 」、高速再生を行う場合には「 $1 < N < \infty$ 」、スキップ再生を行う場合には「 $N=\infty$ 」、スロー逆再生を行う場合には「 $-1 < N < 0$ 」、等倍速逆再生を行う場合には「 $N=-1$ 」、高速逆再生を行う場合には「 $N < -1$ 」を入力する。なお、「 $\infty$ 」は計算機上では十分大きな数値を代用して用いればよい。

#### 【0023】

音声再生制御部 105a は、単位分割部 103a から入力されたテキスト単位データと、入力部 104 から入力された再生条件とに基づいて、単位分割部 103a にて分割した単位毎の再生時間  $l_i$  を算出し、これに基づいて音声データ生成部 101 から入力される音声データをアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0024】

次に、上記構成の話速可変装置の動作について説明する。

ユーザから発せられた音声は、マイクロホン 100 によりアナログ音声信号に変換された後、音声データ生成部 101 に入力され、ここで音声データに変換される。この音声データは、テキストデータ生成部 102 および音声再生制御部 105a に出力される。

#### 【0025】

上記音声データは、テキストデータ生成部 102 にて解析され、テキストデー

タに変換される。図2に、ユーザから発せられた音声をテキストデータに変換した一例を示す。なお、図2において「△」は、所定時間以上の間、無音だったことを示すものであって、テキストデータ生成部102により上記テキストデータに挿入される。

#### 【0026】

次に、単位分割部103aが、上記テキストデータを言語的に意味を持つ単位に分割する。以下の説明では、「形態素」（以下、単位と略称する）で分割した場合について述べる。図2に示すテキストデータを「形態素」単位に分割した場合、図3に示すように分割される。

#### 【0027】

この図において「/」は単位の区切りを示すものである。なお、各単位に対応する数字は、その音声の入力時の再生時間を示し、ここでは総再生時間L0が60秒の例を示している。

#### 【0028】

その後、単位分割部103aは、各単位に、図4に示すような単位の種別に応じた再生情報を設定する。これにより、図5に示すように、各単位に再生情報が対応づけられたテキスト単位データが生成され、これは音声再生制御部105aに出力される。

#### 【0029】

一方、入力部104からは、再生倍率Nと総再生時間Lが入力され、この情報は音声再生制御部105aに通知される。ここでは、入力音声の総再生時間L0が60秒であるのに対して、総再生時間Lとして40秒が指定され、再生倍率Nとして3倍が指定されたものとして説明する。

#### 【0030】

そして、音声再生制御部105aは、単位分割部103aからテキスト単位データが入力されるとともに、入力部104から再生倍率Nと総再生時間Lが入力されると、これらの情報に基づいて、各単位の再生倍率を下記の式(1)にしたがって算出する。

#### 【0031】

## 【数 1】

$$(w_i \text{ の再生倍率}) = \frac{1}{a_i} \times N \times \frac{L'}{L} = \frac{1}{a_i} \times N \times \frac{\sum_i (a_i \cdot l_i)}{\sum_i l_i} \dots (1)$$

## 【0032】

なお、上式(1)において、「i」は各単位の順位、「wi」は単位(形態素)、「li」は各単位の再生時間、「ai」は再生情報に基づく重み、「L0」は入力音声のオリジナルの総再生時間、「L」は重み付け後の総再生時間、「N」は再生倍率を示す。

上式(1)により、この例における形態素の種別(品詞)に応じた再生倍率は、図6に示すような再生倍率となる。

## 【0033】

これにより、音声再生制御部105aは、各単位の再生時間liを図7に示すような値で算出し、この情報を再生制御データとする。そしてこの再生制御データにしたがって、音声データ生成部101から入力された音声データを再生制御し、スピーカ106より拡声出力する。

## 【0034】

以上のように、上記構成の話速可変装置では、従来のようにすべての音声データを一律の速度で再生するのではなく、音声データからテキストデータに変換し、このテキストデータを解析して、例えば形態素などの単位毎に、その重要度に応じて再生速度(再生時間)を可変制御し、全体として所望の再生速度(再生時間)となるようにしている。

## 【0035】

したがって、上記構成の話速可変装置によれば、内容を把握する上で重要な単位については、低速で再生し、重要でないものはより高速で再生することができるので、高速再生を行っても再生される音声を示す内容を理解しやすい。

## 【0036】

次に、この発明の第2の実施形態に係わる話速可変装置について説明する。図8は、その構成を示すものである。



この図に示す話速可変装置は、マイクロホン100、音声データ生成部101、テキストデータ生成部102、単位分割部103b、入力部104、音声再生制御部105b、スピーカ106および記憶部110を備えている。

#### 【0037】

音声データ生成部101は、マイクロホン100から入力されるアナログ音声信号を音声データに変換し、これをテキストデータ生成部102および音声再生制御部105bに出力する。

#### 【0038】

テキストデータ生成部102は、音声認識技術を駆使して、音声データ生成部101から入力された音声データに基づいて、その発声内容を解析してテキストデータに変換する。

#### 【0039】

単位分割部103bは、テキストデータ生成部102で得られたテキストデータを、言語的に意味を持つ単位に分割し、各単位に再生情報を設定し、これをテキスト単位データとして音声再生制御部105bに出力する。

#### 【0040】

ここで、言語的に意味を持つ単位とは、例えば「形態素」、「単語」、「文節」、「連文節」、「文」、「段落」のいずれかを意味する。そして、再生情報とは、分割された単位の種別（例えば品詞）に応じた重み付けの値 $\alpha$ である。

#### 【0041】

入力部104は、例えばオペレータや外部機器から、音声再生の速度に関わる情報として、再生倍率 $N$ と総再生時間 $L$ の入力を受け付けるものである。ここから入力された情報は、後述する音声再生制御部105b内のメモリに記憶される。

#### 【0042】

ここで、等倍速再生を行う場合には「 $N=1$ 」、一時停止する場合には「 $N=0$ 」、スロー再生を行う場合には「 $0<N<1$ 」、高速再生を行う場合には「 $1<N<\infty$ 」、スキップ再生を行う場合には「 $N=\infty$ 」、スロー逆再生を行う場合には「 $-1<N<0$ 」、等倍速逆再生を行う場合には「 $N=-1$ 」、高速逆再生を行う場合には「 $N<-1$ 」を入力する。なお、「 $\infty$ 」は計算機上では十分大きな数値を代用して用いればよい。

**【0043】**

音声再生制御部105bは、音声データ生成部101から入力された音声データと単位分割部103bから入力されたテキスト単位データとを記憶部110に記録する。

**【0044】**

記憶部110は、例えば半導体メモリやハードディスク、光学式の記録媒体などが考えられ、音声再生制御部105bによりデータの読み書きが自在になされる。上記記録媒体は、当該話速可変装置より着脱できることが望ましい。

**【0045】**

そして、音声再生制御部105bは、上記テキスト単位データと、入力部104から入力された再生条件とに基づいて、単位分割部103bにて分割した単位毎の再生時間 $li$ を算出し、これを再生制御データとして記憶部110に記録する。

**【0046】**

そして、音声再生制御部105bは、入力部104を通じたユーザからの要求に応じて、記憶部110に記憶される音声データを、再生制御データに含まれる再生時間 $li$ で再生するようにアナログ音声信号に変換し、スピーカ106より拡声出力する。

**【0047】**

次に、上記構成の話速可変装置の動作について説明する。

ユーザから発せられた音声は、マイクロホン100によりアナログ音声信号に変換された後、音声データ生成部101に入力され、ここで音声データに変換される。この音声データは、テキストデータ生成部102および音声再生制御部105bに出力される。

**【0048】**

上記音声データは、テキストデータ生成部102にて解析され、テキストデータに変換される。このテキストデータの一例を図2に示す。なお、図2において「△」は、所定時間以上の間、無音だったことを示すものであって、テキストデータ生成部102により上記テキストデータに挿入される。

**【0049】**

次に、単位分割部103bが、上記テキストデータを言語的に意味を持つ単位に分割する。以下の説明では、「形態素」単位（以下、単位と略称する）で分割した場合について述べる。「形態素」単位に分割した場合、上記テキストデータは図3に示すように分割される。

**【0050】**

この図において「/」は単位の区切りを示すものである。なお、各単位に対応する数字は、その音声の入力時の再生時間を示し、ここでは総再生時間L0が60秒の例を示している。

**【0051】**

その後、単位分割部103bは、各単位に、図4に示すような単位の種別に応じた再生情報を設定する。これにより、図5に示すように、各単位に再生情報が対応づけられたテキスト単位データが生成され、これは音声再生制御部105bに出力される。

上記音声データと上記テキスト単位データは、音声再生制御部105bによって記憶部110に記録される。

**【0052】**

一方、入力部104からは、再生倍率Nと総再生時間Lが入力され、この情報は音声再生制御部105bに通知される。ここでは、入力音声の総再生時間L0が60秒であるのに対して、総再生時間Lとして40秒が指定され、再生倍率Nとして3倍が指定されたものとして説明する。

**【0053】**

そして、音声再生制御部105bは、入力部104から再生倍率Nと総再生時間Lが入力されると、これらの情報と第1の実施形態で述べた式(1)に基づいて、各単位の再生倍率を算出する。

**【0054】**

そしてさらに、音声再生制御部105bは、図7に示すような各単位の再生時間 $l_i$ を算出し、これを再生制御データとして記憶部110に記録する。

**【0055】**

その後、音声再生制御部105bは、入力部104を通じたユーザからの要求に応じて、記憶部110に記憶される音声データを、再生制御データに含まれる再生時間liで再生するようにアナログ音声信号に変換し、スピーカ106より拡声出力する。

#### 【0056】

以上のように、上記構成の話速可変装置では、従来のようにすべての音声データを一律の速度で再生するのではなく、音声データからテキストデータに変換し、このテキストデータを解析して、例えば形態素などの単位毎に、その重要度に応じて再生速度（再生時間）を可変制御し、全体として所望の再生速度（再生時間）となるようにしている。

#### 【0057】

したがって、上記構成の話速可変装置によれば、内容を把握する上で重要な単位については、低速で再生し、重要でないものはより高速で再生することができるので、高速再生を行っても再生される音声が表示内容を理解しやすい。

#### 【0058】

また、上記構成の話速可変装置では、各単位の再生時間liを算出し、これを音声データ生成部101から入力される音声データとともに記憶部110に記録しておき、ユーザの要求に応じて、各単位に対応する再生時間liで音声データを再生するようにしている。

#### 【0059】

したがって、上記構成の話速可変装置によれば、再生に先立って、音声データとこれに対応する各単位の再生時間liとが、予め対応づけられて生成されるため、ユーザの要求に応じて迅速に再生することができる。

#### 【0060】

また、記憶部110の記録媒体には、単位毎の再生時間liにしたがって音声データを再生することで、音声が表示内容を理解しやすく高速再生できる情報を記録でき、この記録媒体を他の再生機器に読み込ませることで、高速再生が可能である。

#### 【0061】

次に、この発明の第3の実施形態に係わる話速可変装置について説明する。図9は、その構成を示すものである。

この図に示す話速可変装置は、マイクロホン100、音声データ生成部101、テキストデータ生成部102、単位分割部103c、入力部104、音声再生制御部105cおよびスピーカ106を備えている。

#### 【0062】

音声データ生成部101は、マイクロホン100から入力されるアナログ音声信号を音声データに変換し、これをテキストデータ生成部102および音声再生制御部105cに出力する。

#### 【0063】

テキストデータ生成部102は、音声認識技術を駆使して、音声データ生成部101から入力された音声データに基づいて、その発声内容を解析してテキストデータに変換する。

#### 【0064】

単位分割部103cは、テキストデータ生成部102で得られたテキストデータを、言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部105cに出力する。

ここで、言語的に意味を持つ単位とは、例えば「形態素」、「単語」、「文節」、「連文節」、「文」、「段落」のいずれかを意味する。

#### 【0065】

入力部104は、例えばオペレータや外部機器から、音声再生の速度に関わる情報として、再生倍率Nの入力を受け付けるものである。ここから入力された情報は、後述する音声再生制御部105c内のメモリに記憶される。

#### 【0066】

ここで、等倍速再生を行う場合には「 $N=1$ 」、一時停止する場合には「 $N=0$ 」、スロー再生を行う場合には「 $0 < N < 1$ 」、高速再生を行う場合には「 $1 < N < \infty$ 」、スキップ再生を行う場合には「 $N=\infty$ 」、スロー逆再生を行う場合には「 $-1 < N < 0$ 」、等倍速逆再生を行う場合には「 $N=-1$ 」、高速逆再生を行う場合には「 $N < -1$ 」を入力する。なお、「 $\infty$ 」は計算機上では十分大きな数値を代用して用いればよい。

**【0067】**

音声再生制御部 105c は、単位分割部 103c が作成したテキスト単位データと、入力部 104 から入力された再生倍率Nとに基づいて、単位分割部 103c にて分割した単位を選択的に用いて、音声データをアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

**【0068】**

次に、上記構成の話速可変装置の動作について説明する。

ユーザから発せられた音声は、マイクロホン 100 によりアナログ音声信号に変換された後、音声データ生成部 101 に入力され、ここで音声データに変換される。この音声データは、テキストデータ生成部 102 および音声再生制御部 105c に出力される。

**【0069】**

上記音声データは、テキストデータ生成部 102 にて解析され、テキストデータに変換される。図 10 に、ユーザから発せられた音声をテキストデータに変換した一例を示す。

**【0070】**

次に、単位分割部 103c が、上記テキストデータを言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部 105c に出力する。以下の説明では、「文節」で分割した場合について述べる。

**【0071】**

図 10 に示したテキストデータを「文節」単位（以下、単位と略称する）に分割した場合、図 11 に示すように分割される。この図において「|」は単位（文節）の区切りを示すものである。

**【0072】**

一方、入力部 104 からは、再生倍率Nが入力され、この情報は音声再生制御部 105c に通知される。ここでは、再生倍率Nとして2倍が指定されたものとして説明する。

**【0073】**

そして、音声再生制御部 105c は、単位分割部 103c からテキスト単位データが入力されるとともに、入力部 104 から再生倍率Nが入力されると、再生倍率Nが2であることより、上記テキスト単位データに基づいて図 12 に示すように、1つおきの単位を示す再生制御データを生成し、これにしたがって1つおきの単位に対応する音声データをアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0074】

なお、音声再生制御部 105c は、再生倍率Nとして3倍が指定された場合には、3つの連続する単位のうち、1つの単位に対応する音声データを再生し、再生倍率NとしてM倍が指定された場合には、M個の連続する単位のうち、1つの単位に対応する音声データを再生する。

#### 【0075】

以上のように、上記構成の話速可変装置では、従来のように一律の速度ですべての音声データを再生するのではなく、音声データからテキストデータに変換し、このテキストデータを解析して、指定される速度に応じて、例えば文節などの単位を選択的に間引きして再生するようにしている。

#### 【0076】

したがって、上記構成の話速可変装置によれば、間引かれなかった単位の音声は通常速度で再生されることになるので、再生に要する時間を短縮しつつも、再生される音声を示す内容を理解しやすい。

#### 【0077】

次に、この発明の第4の実施形態に係わる話速可変装置について説明する。図 13 は、その構成を示すものである。

この図に示す話速可変装置は、マイクロホン 100、音声データ生成部 101、テキストデータ生成部 102、単位分割部 103d、入力部 104、音声再生制御部 105d、スピーカ 106 および記憶部 110 を備えている。

#### 【0078】

音声データ生成部 101 は、マイクロホン 100 から入力されるアナログ音声信号を音声データに変換し、これをテキストデータ生成部 102 および音声再生

制御部 105d に出力する。

【0079】

テキストデータ生成部 102 は、音声認識技術を駆使して、音声データ生成部 101 から入力された音声データに基づいて、その発声内容を解析してテキストデータに変換する。

【0080】

単位分割部 103d は、テキストデータ生成部 102 で得られたテキストデータを、言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部 105d に出力する。

ここで、言語的に意味を持つ単位とは、例えば「形態素」、「単語」、「文節」、「連文節」、「文」、「段落」のいずれかを意味する。

【0081】

入力部 104 は、例えばオペレータや外部機器から、音声再生の速度に関わる情報として、再生倍率  $N$  の入力を受け付けるものである。ここから入力された情報は、後述する音声再生制御部 105d 内のメモリに記憶される。

【0082】

ここで、等倍速再生を行う場合には「 $N=1$ 」、一時停止する場合には「 $N=0$ 」、スロー再生を行う場合には「 $0 < N < 1$ 」、高速再生を行う場合には「 $1 < N < \infty$ 」、スキップ再生を行う場合には「 $N=\infty$ 」、スロー逆再生を行う場合には「 $-1 < N < 0$ 」、等倍速逆再生を行う場合には「 $N=-1$ 」、高速逆再生を行う場合には「 $N < -1$ 」を入力する。なお、「 $\infty$ 」は計算機上では十分大きな数値を代用して用いればよい。

【0083】

音声再生制御部 105d は、音声データ生成部 101 から入力された音声データと単位分割部 103d から入力されたテキスト単位データとを記憶部 110 に記録する。

【0084】

記憶部 110 は、例えば半導体メモリやハードディスク、光学式の記録媒体などが考えられ、音声再生制御部 105d によりデータの読み書きが自在になされる。上記記録媒体は、当該話速可変装置より着脱できることが望ましい。



## 【0085】

そして、音声再生制御部105dは、上記テキスト単位データと、入力部104から入力された再生倍率Nとに基づいて、単位分割部103dにて分割した単位を選択し、この選択した単位を識別する再生制御データを記憶部110に記録する。

## 【0086】

そして、音声再生制御部105dは、入力部104を通じたユーザからの要求に応じて、記憶部110に記憶される音声データのうち、再生制御データで示される部分についてアナログ音声信号に変換し、スピーカ106より拡声出力する。

## 【0087】

次に、上記構成の話速可変装置の動作について説明する。

ユーザから発せられた音声は、マイクロホン100によりアナログ音声信号に変換された後、音声データ生成部101に入力され、ここで音声データに変換される。この音声データは、テキストデータ生成部102および音声再生制御部105dに出力される。

## 【0088】

上記音声データは、テキストデータ生成部102にて解析され、テキストデータに変換される。図10に、ユーザから発せられた音声をテキストデータに変換した一例を示す。

## 【0089】

次に、単位分割部103dが、上記テキストデータを言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部105dに出力する。

## 【0090】

上記音声データと上記テキスト単位データは、音声再生制御部105dによって記憶部110に記録される。以下の説明では、「文節」で分割した場合について述べる。

## 【0091】

図10に示したテキストデータを「文節」単位（以下、単位と略称する）に分割した場合、図11に示すように分割される。この図において「|」は単位（文節）の区切りを示すものである。

#### 【0092】

一方、入力部104からは、再生倍率Nが入力され、この情報は音声再生制御部105dに通知される。ここでは、再生倍率Nとして2倍が指定されたものとして説明する。

#### 【0093】

そして、音声再生制御部105dは、単位分割部103dからテキスト単位データが入力されるとともに、入力部104から再生倍率Nが入力されると、再生倍率Nが2であることより、上記テキスト単位データに基づいて図12に示すように、1つおきの単位を示す再生制御データを生成し、これを記憶部110に記録する。

#### 【0094】

その後、音声再生制御部105dは、入力部104を通じたユーザからの要求に応じて、記憶部110に記憶される音声データのうち、上記再生制御データで示される単位に対応する文についてアナログ音声信号に変換し、スピーカ106より拡声出力する。

#### 【0095】

以上のように、上記構成の話速可変装置では、従来のように一律の速度ですべての音声データを再生するのではなく、音声データからテキストデータに変換し、このテキストデータを解析して、指定される速度に応じて、例えば文節などの単位を選択的に間引きして再生するようにしている。

#### 【0096】

したがって、上記構成の話速可変装置によれば、間引かれなかった単位の音声は通常速度で再生されることになるので、再生に要する時間を短縮しつつも、再生される音声を示す内容を理解しやすい。

#### 【0097】

また、上記構成の話速可変装置では、ユーザから指定される再生倍率Nに応じ

て、再生すべき単位を示す単位識別情報を音声データとともに記憶部 110 に記録しておき、ユーザの要求に応じて、上記単位識別情報に対応する音声データを再生するようにしている。

#### 【0098】

したがって、上記構成の話速可変装置によれば、再生に先立って、予め再生すべき単位を示す単位識別情報を作成して音声データとともに保存しておくため、ユーザの要求に応じて迅速に再生することができる。

#### 【0099】

また、記憶部 110 の記録媒体には、単位識別情報にしたがって音声データを再生することで、音声が表示内容を理解しやすく高速再生できる情報を記録でき、これによりこの記録媒体を他の再生機器に読み込ませることで、高速再生が可能である。

#### 【0100】

次に、この発明の第 5 の実施形態に係わる話速可変装置について説明する。図 14 は、その構成を示すものである。

この図に示す話速可変装置は、マイクロホン 100、音声データ生成部 101、テキストデータ生成部 102、単位分割部 103e、入力部 104、音声再生制御部 105e およびスピーカ 106 を備えている。

#### 【0101】

音声データ生成部 101 は、マイクロホン 100 から入力されるアナログ音声信号を音声データに変換し、これをテキストデータ生成部 102 および音声再生制御部 105e に出力する。

#### 【0102】

テキストデータ生成部 102 は、音声認識技術を駆使して、音声データ生成部 101 から入力された音声データに基づいて、その発声内容を解析してテキストデータに変換する。

#### 【0103】

単位分割部 103e は、テキストデータ生成部 102 で得られたテキストデータを、言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテ

キスト単位データとして音声再生制御部 105 e に出力する。

ここで、言語的に意味を持つ単位とは、例えば「単語」、「文節」、「連文節」、「文」、「段落」のいずれかを意味する。

#### 【0104】

入力部 104 は、例えばオペレータや外部機器から、音声再生の速度に関わる情報として、キーワードと抽出閾値の入力を受け付けるものである。ここから入力された情報は、後述する音声再生制御部 105 e 内のメモリに記憶される。

#### 【0105】

音声再生制御部 105 e は、単位分割部 103 e が作成したテキスト単位データと、入力部 104 から入力されたキーワードと抽出閾値とに基づいて、単位分割部 103 e にて分割した単位を選択的に用いて、音声データをアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0106】

次に、上記構成の話速可変装置の動作について説明する。

ユーザから発せられた音声は、マイクロホン 100 によりアナログ音声信号に変換された後、音声データ生成部 101 に入力され、ここで音声データに変換される。この音声データは、テキストデータ生成部 102 および音声再生制御部 105 e に出力される。

#### 【0107】

上記音声データは、テキストデータ生成部 102 にて解析され、テキストデータに変換される。図 15 に、ユーザから発せられた音声をテキストデータに変換した一例を示す。

#### 【0108】

次に、単位分割部 103 e が、上記テキストデータを言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部 105 e に出力する。以下の説明では、「文」で分割した場合について述べる。

#### 【0109】

一方、入力部 104 からは、キーワードと抽出閾値が入力され、この情報は音

声再生制御部 105e に通知される。ここでは、キーワードとして、「再生」、「時間」、「制御」の 3 つが入力され、抽出閾値として「2」が入力されたものとして説明する。

#### 【0110】

これに対して、音声再生制御部 105e は、単位分割部 103e からテキスト単位データが入力されるとともに、入力部 104 から上記 3 つのキーワードおよび抽出閾値が入力されると、上記テキスト単位データに基づいて、図 15 に示すように各文から上記キーワードを検出し、これに基づいて、図 16 に示すように 1 つの文に上記キーワードをのべ 2 つ以上含むものを検出する。

そして、音声再生制御部 105e は、上記検出した文の識別情報を示す再生制御データを生成し、この再生制御データに基づいて、上記検出した文に対応する音声データをアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0111】

以上のように、上記構成の話速可変装置では、従来のように一律の速度ですべての音声データを再生するのではなく、音声データからテキストデータに変換し、このテキストデータを解析して、指定されるキーワードを指定された数（抽出閾値）以上含む単位を抽出し、この抽出した単位だけを通常速度で再生するようにしている。

#### 【0112】

したがって、上記構成の話速可変装置によれば、ユーザが必要とする単位についてのみ通常速度で音声再生されることになるので、再生に要する時間を短縮しつつも、再生される音声が表示内容を理解しやすい。

#### 【0113】

次に、この発明の第 6 の実施形態に係わる話速可変装置について説明する。図 17 は、その構成を示すものである。

この図に示す話速可変装置は、マイクロホン 100、音声データ生成部 101、テキストデータ生成部 102、単位分割部 103f、入力部 104、音声再生制御部 105f、スピーカ 106 および記憶部 110 を備えている。

#### 【0114】

音声データ生成部 101 は、マイクロホン 100 から入力されるアナログ音声信号を音声データに変換し、これをテキストデータ生成部 102 および音声再生制御部 105 f に出力する。

#### 【0115】

テキストデータ生成部 102 は、音声認識技術を駆使して、音声データ生成部 101 から入力された音声データに基づいて、その発声内容を解析してテキストデータに変換する。

#### 【0116】

単位分割部 103 f は、テキストデータ生成部 102 で得られたテキストデータを、言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部 105 f に出力する。

ここで、言語的に意味を持つ単位とは、例えば「単語」、「文節」、「連文節」、「文」、「段落」のいずれかを意味する。

#### 【0117】

入力部 104 は、例えばオペレータや外部機器から、音声再生の速度に関わる情報として、キーワードと抽出閾値の入力を受け付けるものである。ここから入力された情報は、後述する音声再生制御部 105 f 内のメモリに記憶される。

#### 【0118】

音声再生制御部 105 f は、音声データ生成部 101 から入力された音声データと単位分割部 103 f から入力されたテキスト単位データとを記憶部 110 に記録する。

#### 【0119】

記憶部 110 は、例えば半導体メモリやハードディスク、光学式の記録媒体などが考えられ、音声再生制御部 105 f によりデータの読み書きが自在になされる。上記記録媒体は、当該話速可変装置より着脱できることが望ましい。

#### 【0120】

そして、音声再生制御部 105 f は、上記テキスト単位データと、入力部 104 から入力されたキーワードと抽出閾値とに基づいて、単位分割部 103 f にて分割した単位を選択し、この選択した単位を示す識別情報を再生制御データとし

て記憶部 110 に記録する。

#### 【0121】

そして、音声再生制御部 105 f は、入力部 104 を通じたユーザからの要求に応じて、記憶部 110 に記憶される音声データのうち、再生制御データで示される部分についてアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0122】

次に、上記構成の話速可変装置の動作について説明する。

ユーザから発せられた音声は、マイクロホン 100 によりアナログ音声信号に変換された後、音声データ生成部 101 に入力され、ここで音声データに変換される。この音声データは、テキストデータ生成部 102 および音声再生制御部 105 f に出力される。

#### 【0123】

上記音声データは、テキストデータ生成部 102 にて解析され、テキストデータに変換される。図 15 に、ユーザから発せられた音声をテキストデータに変換した一例を示す。

#### 【0124】

次に、単位分割部 103 f が、上記テキストデータを言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部 105 f に出力する。

#### 【0125】

上記音声データと上記テキスト単位データは、音声再生制御部 105 f によって記憶部 110 に記録される。以下の説明では、「文」で分割した場合について述べる。

#### 【0126】

一方、入力部 104 からは、キーワードと抽出閾値が入力され、この情報は音声再生制御部 105 f に通知される。ここでは、キーワードとして、「再生」、「時間」、「制御」の 3 つが入力され、抽出閾値として「2」が入力されたものとして説明する。

## 【0127】

これに対して、音声再生制御部105fは、単位分割部103fからテキスト単位データが入力されるとともに、入力部104から上記3つのキーワードおよび抽出閾値が入力されると、上記テキスト単位データに基づいて、図15に示すように各文から上記キーワードを検出し、これに基づいて、図16に示すように1つの文に上記キーワードをのべ2つ以上含むものを検出し、この検出した文を示す識別情報を再生制御データとして記憶部110に記録する。

## 【0128】

その後、音声再生制御部105fは、入力部104を通じたユーザからの要求に応じて、記憶部110に記憶される音声データのうち、上記再生制御データで示される単位に対応する文についてアナログ音声信号に変換し、スピーカ106より拡声出力する。

## 【0129】

以上のように、上記構成の話速可変装置では、従来のように一律の速度ですべての音声データを再生するのではなく、音声データからテキストデータに変換し、このテキストデータを解析して、指定されるキーワードを指定された数（抽出閾値）以上含む単位を抽出し、この抽出した単位だけを通常速度で再生するようにしている。

## 【0130】

したがって、上記構成の話速可変装置によれば、ユーザが必要とする単位についてのみ通常速度で音声再生されることになるので、再生に要する時間を短縮しつつも、再生される音声が表示内容を理解しやすい。

## 【0131】

また、上記構成の話速可変装置では、ユーザの要求に応じた、再生すべき単位を示す単位識別情報を音声データとともに記憶部110に記録しておき、ユーザの要求に応じて、上記単位識別情報に対応する音声データを再生するようにしている。

## 【0132】

したがって、上記構成の話速可変装置によれば、再生に先立って、予め再生す



べき単位を示す単位識別情報を作成して音声データとともに保存しておくため、ユーザの要求に応じて迅速に再生することができる。

#### 【0133】

また、記憶部110の記録媒体には、単位識別情報にしたがって音声データを再生することで、音声が表示内容を理解しやすく高速再生できる情報を記録でき、これによりこの記録媒体を他の再生機器に読み込ませることで、高速再生が可能である。

#### 【0134】

次に、この発明の第7の実施形態に係わる話速可変装置について説明する。図18は、その構成を示すものである。

この図に示す話速可変装置は、マイクロホン100、音声データ生成部101、テキストデータ生成部102、単位分割部103g、入力部104、音声再生制御部105g、スピーカ106および記憶部120を備えている。

#### 【0135】

音声データ生成部101は、マイクロホン100から入力されるアナログ音声信号を音声データに変換し、これをテキストデータ生成部102および音声再生制御部105gに出力する。

#### 【0136】

テキストデータ生成部102は、音声認識技術を駆使して、音声データ生成部101から入力された音声データに基づいて、その発声内容を解析してテキストデータに変換する。

#### 【0137】

単位分割部103gは、テキストデータ生成部102で得られたテキストデータを、言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部105gに出力する。

ここで、言語的に意味を持つ単位とは、例えば「形態素」、「単語」、「文」、「段落」のいずれかを意味する。

#### 【0138】

入力部104は、例えばオペレータや外部機器から、音声再生の速度に関わる

情報として、再生倍率 $N$ の入力を受け付けるものである。ここから入力された情報は、後述する音声再生制御部 105 g 内のメモリに記憶される。

#### 【0139】

ここで、等倍速再生を行う場合には「 $N=1$ 」、一時停止する場合には「 $N=0$ 」、スロー再生を行う場合には「 $0 < N < 1$ 」、高速再生を行う場合には「 $1 < N < \infty$ 」、スキップ再生を行う場合には「 $N=\infty$ 」、スロー逆再生を行う場合には「 $-1 < N < 0$ 」、等倍速逆再生を行う場合には「 $N=-1$ 」、高速逆再生を行う場合には「 $N < -1$ 」を入力する。なお、「 $\infty$ 」は計算機上では十分大きな数値を代用して用いればよい。

#### 【0140】

音声再生制御部 105 g は、単位分割部 103 g が作成したテキスト単位データと、入力部 104 から入力された再生倍率 $N$ と、後述する記憶部 120 に記憶される言語的優先度情報とに基づいて、単位分割部 103 g にて分割した単位を選択的に用いて、音声データをアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0141】

記憶部 120 は、例えば図 19 に示すような、多数の接続詞をその重要度に応じて分類した情報を、上記言語的優先度情報として記憶するものである。

#### 【0142】

次に、上記構成の話速可変装置の動作について説明する。

ユーザから発せられた音声は、マイクロホン 100 によりアナログ音声信号に変換された後、音声データ生成部 101 に入力され、ここで音声データに変換される。この音声データは、テキストデータ生成部 102 および音声再生制御部 105 g に出力される。

#### 【0143】

上記音声データは、テキストデータ生成部 102 にて解析され、テキストデータに変換される。

次に、単位分割部 103 g が、上記テキストデータを言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部 105 g に出力する。以下の説明では、「文」で分割した場合について

述べる。

#### 【0144】

図20に、ユーザから発せられた音声をテキストデータに変換した一例を示す。この図では、各文(1)～(8)を省略して示し、先頭に置かれた接続詞のみを表記している。

#### 【0145】

一方、入力部104からは、再生倍率Nが入力され、この情報は音声再生制御部105gに通知される。ここでは、再生倍率Nとして2倍が指定されたものとして説明する。

#### 【0146】

これに対して、音声再生制御部105gは、単位分割部103gからテキスト単位データが入力されるとともに、入力部104から再生倍率Nが入力されると、各文(1)～(8)の再生時間を検出するとともに、すべての文を再生するのに要する総再生時間を検出する。ここでは、総再生時間は80秒であったとする。

#### 【0147】

次に、音声再生制御部105gは、再生倍率Nが2で、総再生時間が80秒であることより、実再生時間を40秒以内と決定する。

そして、音声再生制御部105gは、記憶部120に記憶される言語的優先度情報を参照し、テキスト単位データから各文(1)～(8)のうち、重要度の高い接続詞が先頭に置かれるものを順に検出し、このうち合計の再生時間が実再生時間40秒にできるだけ近く、かつ40秒以内となるように、文を選択する。

#### 【0148】

この選択した文の識別情報は、音声再生制御部105hにより再生制御データとして生成される。なお、図20に示す例では、図21に示すように、文(1)、(3)、(6)、(8)の合わせて40秒分が選択されることになる。

#### 【0149】

そして、音声再生制御部105gは、上記再生制御データに基づいて、記憶部110に記憶される音声データのうち、上記選択した文(1)、(3)、(6)

、(8)に対応する音声データをアナログ音声信号に変換し、スピーカ106より拡声出力する。

【0150】

以上のように、上記構成の話速可変装置では、従来のように一律の速度ですべての音声データを再生するのではなく、音声データからテキストデータに変換し、このテキストデータを解析して、予め高い優先度が設定された文字列を含む単位を抽出し、この抽出した単位だけを通常速度で再生するようにしている。

【0151】

したがって、上記構成の話速可変装置によれば、予め高い優先度が設定された文字列を含む単位の音声優先的に通常速度で再生されることになるので、再生に要する時間を短縮しつつも、再生される音声を示す内容を理解しやすい。

【0152】

また、ユーザから指定される再生倍率Nに応じた再生時間内に収まるように、優先度の高い単位が選択されるため、再生される音声は、ユーザが指定する時間の範囲内で行うことができる。

【0153】

尚、この発明は上記実施の形態に限定されるものではない。

例えば、上記実施の形態では、記憶部120に記憶される言語的優先度情報として接続詞を例に挙げ、この接続詞の優先度で各文の優先度を決定するものとして説明した。

【0154】

しかしこれに限定されるものではなく、これに代わって例えば、各単位について、単語数、文字数、名詞や動詞といった自立語の数、自立語の割合などを求めて、この結果に応じて各単位の優先度を求めるようにしてもよい。

【0155】

また、上記実施の形態では、ユーザから指定される再生倍率Nに応じて再生時間が決定されるものとして説明したが、再生倍率Nが高い値に設定された場合には、再生される文(単位)が少なくなってしまう虞がある。

【0156】

これに対して例えば、再生される単位の数が予め設定した閾値未満となった場合には、音声再生制御部 105 g が、再生倍率 N を低倍率化し、その倍率で決定される再生時間内で再生可能な優先度の高い単位を選択することで、再生される単位の数を増やして再生するようにしてもいい。

#### 【0157】

次に、この発明の第 8 の実施形態に係わる話速可変装置について説明する。図 22 は、その構成を示すものである。

この図に示す話速可変装置は、マイクロホン 100、音声データ生成部 101、テキストデータ生成部 102、単位分割部 103 h、入力部 104、音声再生制御部 105 h、スピーカ 106、記憶部 110 および記憶部 120 を備えている。

#### 【0158】

音声データ生成部 101 は、マイクロホン 100 から入力されるアナログ音声信号を音声データに変換し、これをテキストデータ生成部 102 および音声再生制御部 105 h に出力する。

#### 【0159】

テキストデータ生成部 102 は、音声認識技術を駆使して、音声データ生成部 101 から入力された音声データに基づいて、その発声内容を解析してテキストデータに変換する。

#### 【0160】

単位分割部 103 h は、テキストデータ生成部 102 で得られたテキストデータを、言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部 105 h に出力する。

ここで、言語的に意味を持つ単位とは、例えば「形態素」、「単語」、「文」、「段落」のいずれかを意味する。

#### 【0161】

入力部 104 は、例えばオペレータや外部機器から、音声再生の速度に関わる情報として、再生倍率 N の入力を受け付けるものである。ここから入力された情報は、後述する音声再生制御部 105 h 内のメモリに記憶される。

## 【0162】

ここで、等倍速再生を行う場合には「 $N=1$ 」、一時停止する場合には「 $N=0$ 」、スロー再生を行う場合には「 $0 < N < 1$ 」、高速再生を行う場合には「 $1 < N < \infty$ 」、スキップ再生を行う場合には「 $N=\infty$ 」、スロー逆再生を行う場合には「 $-1 < N < 0$ 」、等倍速逆再生を行う場合には「 $N=-1$ 」、高速逆再生を行う場合には「 $N < -1$ 」を入力する。なお、「 $\infty$ 」は計算機上では十分大きな数値を代用して用いればよい。

## 【0163】

音声再生制御部 105h は、音声データ生成部 101 から入力された音声データと単位分割部 103h から入力されたテキスト単位データとを記憶部 110 に記録する。

## 【0164】

記憶部 110 は、例えば半導体メモリやハードディスク、光学式の記録媒体などが考えられ、音声再生制御部 105h によりデータの読み書きが自在になされる。上記記録媒体は、当該話速可変装置より着脱できることが望ましい。

## 【0165】

そして、音声再生制御部 105h は、上記テキスト単位データと、入力部 104 から入力された再生倍率  $N$  と、後述する記憶部 120 に記憶される言語的優先度情報とに基づいて、再生すべき単位を選択し、この選択した単位を示す識別情報を再生制御データとして記憶部 110 に記録する。

## 【0166】

そして、音声再生制御部 105h は、入力部 104 を通じたユーザからの要求に応じて、記憶部 110 に記憶される音声データのうち、再生制御データで示される部分についてアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

## 【0167】

記憶部 120 は、例えば図 19 に示すような、多数の接続詞をその重要度に応じて分類した情報を、上記言語的優先度情報として記憶するものである。

## 【0168】

次に、上記構成の話速可変装置の動作について説明する。

ユーザから発せられた音声は、マイクロホン100によりアナログ音声信号に変換された後、音声データ生成部101に入力され、ここで音声データに変換される。この音声データは、テキストデータ生成部102および音声再生制御部105hに出力される。

#### 【0169】

上記音声データは、テキストデータ生成部102にて解析され、テキストデータに変換される。

次に、単位分割部103hが、上記テキストデータを言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部105hに出力する。

#### 【0170】

上記音声データと上記テキスト単位データは、音声再生制御部105jによって記憶部110に記録される。以下の説明では、上記テキストデータを「文」で分割した場合を例に挙げて説明する。

#### 【0171】

図20に、ユーザから発せられた音声をテキストデータに変換した一例を示す。この図では、各文(1)～(8)を省略して示し、先頭に置かれた接続詞のみを表記している。

#### 【0172】

一方、入力部104からは、再生倍率Nが入力され、この情報は音声再生制御部105gに通知される。ここでは、再生倍率Nとして2倍が指定されたものとして説明する。

#### 【0173】

これに対して、音声再生制御部105hは、単位分割部103hからテキスト単位データが入力されるとともに、入力部104から再生倍率Nが入力されると、各文(1)～(8)の再生時間を検出するとともに、すべての文を再生するのに要する総再生時間を検出する。ここでは、総再生時間は80秒であったとする。

#### 【0174】

次に、音声再生制御部 105h は、再生倍率Nが2で、総再生時間が80秒であることより、実再生時間を40秒以内と決定する。

そして、音声再生制御部 105h は、記憶部 120 に記憶される言語的優先度情報を参照し、テキスト単位データから各文(1)～(8)のうち、重要度の高い接続詞が先頭に置かれるものを順に検出し、このうち合計の再生時間が実再生時間40秒にできるだけ近く、かつ40秒以内となるように、文を選択する。

#### 【0175】

なお、図20に示す例では、図21に示すように、文(1)、(3)、(6)、(8)の合わせて40秒分が選択されることになる。この選択した文の識別情報は、音声再生制御部 105h により再生制御データとして記憶部 110 に記録される。

#### 【0176】

その後、音声再生制御部 105h は、入力部 104 を通じたユーザからの要求に応じて、上記再生制御データに基づいて、記憶部 110 に記憶される音声データのうち、上記選択した文(1)、(3)、(6)、(8)に対応する音声データをアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0177】

以上のように、上記構成の話速可変装置では、従来のように一律の速度ですべての音声データを再生するのではなく、音声データからテキストデータに変換し、このテキストデータを解析して、予め高い優先度が設定された文字列を含む単位を抽出し、この抽出した単位だけを通常速度で再生するようにしている。

#### 【0178】

したがって、上記構成の話速可変装置によれば、予め高い優先度が設定された文字列を含む単位の音声優先的に通常速度で再生されることになるので、再生に要する時間を短縮しつつも、再生される音声を示す内容を理解しやすい。

#### 【0179】

また、ユーザから指定される再生倍率Nに応じた再生時間内に収まるように、優先度の高い単位が選択されるため、再生される音声は、ユーザが指定する時間の範囲内で行うことができる。



**【0180】**

また、上記構成の話速可変装置では、再生すべき単位を示す単位識別情報を音声データとともに記憶部110に記録しておき、ユーザの要求に応じて、上記単位識別情報に対応する音声データを再生するようにしている。

**【0181】**

したがって、上記構成の話速可変装置によれば、再生に先立って、予め再生すべき単位を示す単位識別情報を作成して音声データとともに保存しておくため、ユーザの要求に応じて迅速に再生することができる。

**【0182】**

また、記憶部110の記録媒体には、単位識別情報にしたがって音声データを再生することで、音声が表示内容を理解しやすく高速再生できる情報を記録でき、これによりこの記録媒体を他の再生機器に読み込ませることで、高速再生が可能である。

**【0183】**

次に、この発明の第9の実施形態に係わる話速可変装置について説明する。図23は、その構成を示すものである。

この図に示す話速可変装置は、マイクロホン100、音声データ生成部101、テキストデータ生成部102、単位分割部103i、入力部104、音声再生制御部105i、スピーカ106および記憶部120を備えている。

**【0184】**

音声データ生成部101は、マイクロホン100から入力されるアナログ音声信号を音声データに変換し、これをテキストデータ生成部102および音声再生制御部105iに出力する。

**【0185】**

テキストデータ生成部102は、音声認識技術を駆使して、音声データ生成部101から入力された音声データに基づいて、その発声内容を解析してテキストデータに変換する。

**【0186】**

単位分割部103iは、テキストデータ生成部102で得られたテキストデー

タを、言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部 105 i に出力する。

ここで、言語的に意味を持つ単位とは、例えば「形態素」、「単語」、「文」、「文節」、「段落」のいずれかを意味する。

#### 【0187】

入力部 104 は、例えばオペレータや外部機器から、音声再生の速度に関わる情報として、再生倍率  $N$  の入力を受け付けるものである。ここから入力された情報は、後述する音声再生制御部 105 i 内のメモリに記憶される。

#### 【0188】

ここで、等倍速再生を行う場合には「 $N=1$ 」、一時停止する場合には「 $N=0$ 」、スロー再生を行う場合には「 $0 < N < 1$ 」、高速再生を行う場合には「 $1 < N < \infty$ 」、スキップ再生を行う場合には「 $N=\infty$ 」、スロー逆再生を行う場合には「 $-1 < N < 0$ 」、等倍速逆再生を行う場合には「 $N=-1$ 」、高速逆再生を行う場合には「 $N < -1$ 」を入力する。なお、「 $\infty$ 」は計算機上では十分大きな数値を代用して用いればよい。

#### 【0189】

音声再生制御部 105 i は、単位分割部 103 i が作成したテキスト単位データと、入力部 104 から入力された再生倍率  $N$  と、後述する記憶部 120 に記憶される統計的優先度情報とに基づいて、単位分割部 103 g にて分割した各単位の再生速度を決定し、この決定結果（再生制御データ）に基づいて音声データをアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0190】

記憶部 120 は、上記統計的優先度情報を記憶するものである。統計的優先度情報は、予め新聞や雑誌などの多種多様な活字媒体に基づいて収集されたもので、上記単位間が結合する確率を示すものである。

#### 【0191】

したがって、この確率は、世間一般でよく使用される表現は値が大きくなり、あまり使われない一般的でない表現は値が小さくなる。このことから、上記確率が大きい表現はよく使われる表現のため、正確に再生しなくても再生音の雰囲気や前後関係から容易に推測できると考えられ、一方、条件付き確率値が小さい表

現は特異な表現であるため前後関係等からその表現を推測することが難しいと考えられる。

#### 【0192】

次に、上記構成の話速可変装置の動作について説明する。

ユーザから発せられた音声は、マイクロホン100によりアナログ音声信号に変換された後、音声データ生成部101に入力され、ここで音声データに変換される。この音声データは、テキストデータ生成部102および音声再生制御部105jに出力される。

#### 【0193】

上記音声データは、テキストデータ生成部102にて解析され、テキストデータに変換される。図24に、ユーザから発せられた音声をテキストデータに変換した一例を示す。

#### 【0194】

次に、単位分割部103jが、上記テキストデータを言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部105jに出力する。なお、以下の説明では、上記テキストデータを「文節」で分割した場合を例に挙げて説明する。

#### 【0195】

一方、入力部104からは、再生倍率Nが入力され、この情報は音声再生制御部105gに通知される。ここでは、再生倍率Nとして2倍が指定されたものとして説明する。

#### 【0196】

これに対して、音声再生制御部105jは、入力部104から再生倍率Nが入力されると、上記音声データとテキスト単位データに基づいて、各文節の再生時間を検出する。ここでは、各文節の再生時間の合計が52秒であったとする。

#### 【0197】

次に、音声再生制御部105jは、実際の再生に要する実再生時間を算出する。この例では、再生倍率Nが2で、再生時間が52秒であることより、実再生時間を26秒以内と算出する。

## 【0198】

そして、音声再生制御部105jは、上記テキスト単位データに基づいて、図24に示すようなテキストデータを、図25に示すような連続する2つの単位の組み合わせ(1)～(23)に分類した後、記憶部120に記憶される統計的優先度情報を参照して、各組み合わせの(1)～(23)の確率値を求める。

## 【0199】

そして、音声再生制御部105iは、上記組み合わせの(1)～(23)のうち、予め設定された確率閾値以下のものについては低速(通常速度)再生するものとし、一方、確率閾値より大きいものについては高速(倍速)再生するものとする。

## 【0200】

ここで、上記確率閾値が「0.01」に設定されていたものとする、組み合わせ(2)、(8)、(9)、(13)、(15)、(19)、(20)、(22)が低速再生の対象となり、一方、(1)、(3)、(4)、(5)、(6)、(7)、(10)、(11)、(12)、(14)、(16)、(17)、(18)、(21)、(23)については高速再生の対象となる。

## 【0201】

そして次に音声再生制御部105iは、低速再生対象の組み合わせ(2)、(8)、(9)、(13)、(15)、(19)、(20)、(22)に関わる単位を通常再生するのに要する時間が合計25秒であることを検出する。

## 【0202】

そして音声再生制御部105iは、これより実再生時間26秒のうち残り1秒間で、高速再生対象の(1)、(3)、(4)、(5)、(6)、(7)、(10)、(11)、(12)、(14)、(16)、(17)、(18)、(21)、(23)に関わる単位を再生する必要があることを検出し、これよりこれらについては27倍速再生することを決定する。

## 【0203】

そして、音声再生制御部105jは、このような単位毎の再生速度に関する情報を再生制御データとして生成し、この再生制御データにしたがって、音声デー

タ生成部 101 から入力された音声データのうち、組み合わせ (2)、(8)、(9)、(13)、(15)、(19)、(20)、(22) に関わる単位に対応する音声データを通常再生するようにアナログ音声信号に変換し、また (1)、(3)、(4)、(5)、(6)、(7)、(10)、(11)、(12)、(14)、(16)、(17)、(18)、(21)、(23) に関わる単位に対応する音声データを 27 倍速再生するようにアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0204】

これにより、27 倍速再生される部分については人間がほとんど内容を聞き取れないため、図 26 に示すような通常再生する部分についてのみ音声として認識される。

#### 【0205】

以上のように、上記構成の話速可変装置では、従来のように一律の速度ですべての音声データを再生するのではなく、音声データからテキストデータに変換し、このテキストデータを解析して、あまり一般的でない表現を優先的に通常速度で再生するようにしている。

#### 【0206】

したがって、上記構成の話速可変装置によれば、前後の内容などから内容を推測しやすい一般的な表現については高速再生し、一方、前後の内容などから内容を推測しにくい一般的でない表現については低速再生するようにしているので、再生に要する時間を短縮しつつも、再生される音声が示す内容を理解しやすい。

#### 【0207】

また、ユーザから指定される再生倍率  $N$  に応じた再生時間内に収まるように、優先度の高い単位が選択されるため、再生される音声は、ユーザが指定する時間の範囲内で行うことができる。

#### 【0208】

尚、この発明は上記実施の形態に限定されるものではない。

例えば、上記実施の形態では、音声再生制御部 105 i が図 25 に示すような連続する 2 つの単位の組み合わせ (1) ~ (23) に分類するようにしたが、連

続する 3 つ以上の単位の組み合わせで分類するようにしてもよい。

#### 【0209】

また確率閾値には予め設定された値を用いるようにしたが、これに代わって例えば、まず予め設定された値を用い、その場合に低速再生する単位の合計時間が、入力部 104 から入力された再生倍率  $N$  に基づく実再生時間を超えるような場合には、音声再生制御部 105 i が上記確率閾値を可変して、上記実再生時間内に収まるように調整するようにしてもよい。

#### 【0210】

次に、この発明の第 10 の実施形態に係わる話速可変装置について説明する。図 27 は、その構成を示すものである。

この図に示す話速可変装置は、マイクロホン 100、音声データ生成部 101、テキストデータ生成部 102、単位分割部 103 j、入力部 104、音声再生制御部 105 j、スピーカ 106、記憶部 110 および記憶部 120 を備えている。

#### 【0211】

音声データ生成部 101 は、マイクロホン 100 から入力されるアナログ音声信号を音声データに変換し、これをテキストデータ生成部 102 および音声再生制御部 105 j に出力する。

#### 【0212】

テキストデータ生成部 102 は、音声認識技術を駆使して、音声データ生成部 101 から入力された音声データに基づいて、その発声内容を解析してテキストデータに変換する。

#### 【0213】

単位分割部 103 j は、テキストデータ生成部 102 で得られたテキストデータを、言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部 105 j に出力する。

ここで、言語的に意味を持つ単位とは、例えば「形態素」、「単語」、「文」、「文節」、「段落」のいずれかを意味する。

#### 【0214】

入力部 104 は、例えばオペレータや外部機器から、音声再生の速度に関わる情報として、再生倍率  $N$  の入力を受け付けるものである。ここから入力された情報は、後述する音声再生制御部 105 j 内のメモリに記憶される。

#### 【0215】

ここで、等倍速再生を行う場合には「 $N=1$ 」、一時停止する場合には「 $N=0$ 」、スロー再生を行う場合には「 $0 < N < 1$ 」、高速再生を行う場合には「 $1 < N < \infty$ 」、スキップ再生を行う場合には「 $N=\infty$ 」、スロー逆再生を行う場合には「 $-1 < N < 0$ 」、等倍速逆再生を行う場合には「 $N=-1$ 」、高速逆再生を行う場合には「 $N < -1$ 」を入力する。なお、「 $\infty$ 」は計算機上では十分大きな数値を代用して用いればよい。

#### 【0216】

音声再生制御部 105 j は、音声データ生成部 101 から入力された音声データと単位分割部 103 j から入力されたテキスト単位データとを記憶部 110 に記録する。

#### 【0217】

記憶部 110 は、例えば半導体メモリやハードディスク、光学式の記録媒体などが考えられ、音声再生制御部 105 j によりデータの読み書きが自在になされる。上記記録媒体は、当該話速可変装置より着脱できることが望ましい。

#### 【0218】

そして、音声再生制御部 105 j は、上記テキスト単位データと、入力部 104 から入力された再生倍率  $N$  と、後述する記憶部 120 に記憶される統計的優先度情報とに基づいて、単位分割部 103 g にて分割した各単位の再生速度を決定し、この決定結果を再生制御データとして記憶部 110 に記録する。

#### 【0219】

そして、音声再生制御部 105 j は、入力部 104 を通じたユーザからの要求に応じて、記憶部 110 に記憶される再生制御データにしたがって記憶部 110 に記憶される音声データをアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0220】

記憶部 120 は、上記統計的優先度情報を記憶するものである。統計的優先度

情報は、予め新聞や雑誌などの多種多様な活字媒体に基づいて収集されたもので、上記単位間が結合する確率を示すものである。

#### 【0221】

したがって、この確率は、世間一般でよく使用される表現は値が大きくなり、あまり使われない一般的でない表現は値が小さくなる。このことから、上記確率が大きい表現はよく使われる表現のため、正確に再生しなくても再生音の雰囲気や前後関係から容易に推測できると考えられ、一方、条件付き確率値が小さい表現は特異な表現であるため前後関係等からその表現を推測することが難しいと考えられる。

#### 【0222】

次に、上記構成の話速可変装置の動作について説明する。

ユーザから発せられた音声は、マイクロホン100によりアナログ音声信号に変換された後、音声データ生成部101に入力され、ここで音声データに変換される。この音声データは、テキストデータ生成部102および音声再生制御部105jに出力される。

#### 【0223】

上記音声データは、テキストデータ生成部102にて解析され、テキストデータに変換される。図24に、ユーザから発せられた音声をテキストデータに変換した一例を示す。

#### 【0224】

次に、単位分割部103jが、上記テキストデータを言語的に意味を持つ単位に分割し、この分割した単位を識別する情報をテキスト単位データとして音声再生制御部105jに出力する。

#### 【0225】

上記音声データと上記テキスト単位データは、音声再生制御部105jによって記憶部110に記録される。なお、以下の説明では、上記テキストデータを「文節」で分割した場合を例に挙げて説明する。

#### 【0226】

一方、入力部104からは、再生倍率Nが入力され、この情報は音声再生制御



部 105 g に通知される。ここでは、再生倍率  $N$  として 2 倍が指定されたものとして説明する。

#### 【0227】

これに対して、音声再生制御部 105 j は、入力部 104 から再生倍率  $N$  が入力されると、記憶部 110 に記録される音声データとテキスト単位データに基づいて、各文節の再生時間を検出する。ここでは、各文節の再生時間の合計が 52 秒であったとする。

#### 【0228】

次に、音声再生制御部 105 j は、実際の再生に要する実再生時間を算出する。この例では、再生倍率  $N$  が 2 で、再生時間が 52 秒であることより、実再生時間を 26 秒以内と算出する。

#### 【0229】

そして、音声再生制御部 105 j は、記憶部 110 に記録されるテキスト単位データに基づいて、図 24 に示すようなテキストデータを、図 25 に示すような連続する 2 つの単位の組み合わせ (1) ~ (23) に分類した後、記憶部 120 に記憶される統計的優先度情報を参照して、各組み合わせの (1) ~ (23) の確率値を求める。

#### 【0230】

そして、音声再生制御部 105 i は、上記組み合わせの (1) ~ (23) のうち、予め設定された確率閾値以下のものについては低速（通常速度）再生するものとし、一方、確率閾値より大きいものについては高速（倍速）再生するものとする。

#### 【0231】

ここで、上記確率閾値が「0.01」に設定されていたものとする、組み合わせ (2)、(8)、(9)、(13)、(15)、(19)、(20)、(22) が低速再生の対象となり、一方、(1)、(3)、(4)、(5)、(6)、(7)、(10)、(11)、(12)、(14)、(16)、(17)、(18)、(21)、(23) についてが高速再生の対象となる。

#### 【0232】

そして次に音声再生制御部 105 i は、低速再生対象の組み合わせ (2)、(8)、(9)、(13)、(15)、(19)、(20)、(22) に関わる単位を通常再生するのに要する時間が合計 25 秒であることを検出し、これより実再生時間 26 秒のうち残り 1 秒間で (1)、(3)、(4)、(5)、(6)、(7)、(10)、(11)、(12)、(14)、(16)、(17)、(18)、(21)、(23) に関わる単位を再生する必要があることを検出し、これよりこれらについては 27 倍速再生することを決定する。

#### 【0233】

そして、このような単位毎の再生速度に関する情報を再生制御データとして記憶部 110 に記録する。

そしてやがて、音声再生制御部 105 j は、入力部 104 を通じたユーザからの要求があると、記憶部 110 に記録される再生制御データにしたがって、音声データ生成部 101 から入力された音声データのうち、組み合わせ (2)、(8)、(9)、(13)、(15)、(19)、(20)、(22) に関わる単位に対応する音声データを通常再生するようにアナログ音声信号に変換し、また (1)、(3)、(4)、(5)、(6)、(7)、(10)、(11)、(12)、(14)、(16)、(17)、(18)、(21)、(23) に関わる単位に対応する音声データを 27 倍速再生するようにアナログ音声信号に変換し、スピーカ 106 より拡声出力する。

#### 【0234】

これにより、27 倍速再生される部分については人間がほとんど内容を聞き取れないため、図 26 に示すような通常再生する部分についてのみ音声として認識される。

#### 【0235】

以上のように、上記構成の話速可変装置では、従来のように一律の速度ですべての音声データを再生するのではなく、音声データからテキストデータに変換し、このテキストデータを解析して、あまり一般的でない表現を優先的に通常速度で再生するようにしている。

#### 【0236】

したがって、上記構成の話速可変装置によれば、前後の内容などから内容を推測しやすい一般的な表現については高速再生し、一方、前後の内容などから内容を推測しにくい一般的でない表現については低速再生するようにしているので、再生に要する時間を短縮しつつも、再生される音声を示す内容を理解しやすい。

#### 【0237】

また、ユーザから指定される再生倍率Nに応じた再生時間内に収まるように、優先度の高い単位が選択されるため、再生される音声は、ユーザが指定する時間の範囲内で行うことができる。

#### 【0238】

また、上記構成の話速可変装置では、各単位の再生速度を示す情報を再生制御データとして音声データとともに記憶部110に記録しておき、ユーザの要求に応じて、上記再生制御データに基づいて音声データを再生するようにしている。

#### 【0239】

したがって、上記構成の話速可変装置によれば、再生に先立って、予め各単位の再生速度を示す情報を音声データとともに保存しておくため、ユーザの要求に応じて迅速に再生することができる。

#### 【0240】

また、記憶部110の記録媒体には、再生制御データに基づいて音声データを再生することで、音声を示す内容を理解しやすく高速再生できる情報を記録でき、これによりこの記録媒体を他の再生機器に読み込ませることで、高速再生が可能である。

#### 【0241】

尚、この発明は、以上の実施の形態に限定されるものではない。

例えば、上記実施の形態では、テキストデータ生成部102が、音声認識技術を駆使して、音声データの発声内容を解析してテキストデータに変換するようにした。

#### 【0242】

これに代わって例えば、テキストデータ生成部102を、オペレータに対して音声データに基づく音声を拡声する拡声装置と、この拡声装置から拡声される音

声を聴いたオペレータがそれを認識して、キーボードなどの入力装置からテキスト入力を行って、これを受け付けてテキストデータを出力するテキスト入力装置とから構成するようにしてもよい。

その他、この発明の要旨を逸脱しない範囲で種々の変形を施しても同様に実施可能であることはいうまでもない。

#### 【0243】

##### 【発明の効果】

以上述べたように、この発明では、音声信号から生成した音声データをテキストデータに変換し、このテキストデータを言語的に意味を持つ単位に分割する分割情報を生成し、音声の再生制御に関わる再生情報と上記分割情報とに基づいて、音声データから上記単位で音声信号を生成するようにしている。

#### 【0244】

したがって、この発明によれば、言語的に意味を持つ単位で音声信号を生成できるので、上記単位毎の制御により再生時間を短縮しても、再生される音声が表示内容を理解しやすい話速可変装置及び話速変換方法を提供できる。

##### 【図面の簡単な説明】

【図1】 この発明に係わる話速可変装置の第1の実施の形態の構成を示す回路ブロック図。

【図2】 図1に示した話速可変装置が、ユーザから入力される音声をテキストデータに変換した一例を示す。

【図3】 図1に示した話速可変装置が生成したテキストデータを単位毎に分割する動作を説明するための図。

【図4】 図1に示した話速可変装置によって品詞毎に与えられる再生情報の一例を示す図。

【図5】 図1に示した話速可変装置によって再生情報が対応づけられたテキストデータの一例を示す図。

【図6】 図1に示した話速可変装置によって品詞毎に設定される再生速度の一例を示す図。

【図7】 図1に示した話速可変装置によって算出された再生時間とテキス

トデータの対応関係の一例を示す図。

【図 8】 この発明に係わる話速可変装置の第 2 の実施の形態の構成を示す回路ブロック図。

【図 9】 この発明に係わる話速可変装置の第 3 の実施の形態の構成を示す回路ブロック図。

【図 10】 図 9 に示した話速可変装置が、ユーザから入力される音声テキストデータを変換した一例を示す。

【図 11】 図 9 に示した話速可変装置が生成したテキストデータを単位毎に分割する動作を説明するための図。

【図 12】 図 9 に示した話速可変装置が再生する音声を選択する動作を説明するための図。

【図 13】 この発明に係わる話速可変装置の第 4 の実施の形態の構成を示す回路ブロック図。

【図 14】 この発明に係わる話速可変装置の第 5 の実施の形態の構成を示す回路ブロック図。

【図 15】 図 14 に示した話速可変装置がキーワードを含む単位を選択する動作を説明するための図。

【図 16】 図 14 に示した話速可変装置が再生するために選択した単位の一例を示す図。

【図 17】 この発明に係わる話速可変装置の第 6 の実施の形態の構成を示す回路ブロック図。

【図 18】 この発明に係わる話速可変装置の第 7 の実施の形態の構成を示す回路ブロック図。

【図 19】 図 18 に示した話速可変装置が着目するキーワードの優先度の一例を示す図。

【図 20】 図 18 に示した話速可変装置が検出した各単位の再生に要する時間の一例を示す図。

【図 21】 図 18 に示した話速可変装置が再生するために選択した単位の一例を示す図。

【図 2 2】 この発明に係わる話速可変装置の第 8 の実施の形態の構成を示す回路ブロック図。

【図 2 3】 この発明に係わる話速可変装置の第 9 の実施の形態の構成を示す回路ブロック図。

【図 2 4】 図 2 3 に示した話速可変装置が、ユーザから入力される音声をテキストデータに変換した一例を示す。

【図 2 5】 図 2 3 に示した話速可変装置が、連続する単位の組み合わせが発生する確率の分析結果の一例を示す。

【図 2 6】 図 2 3 に示した話速可変装置が再生するために選択した単位の一例を示す図。

【図 2 7】 この発明に係わる話速可変装置の第 10 の実施の形態の構成を示す回路ブロック図。

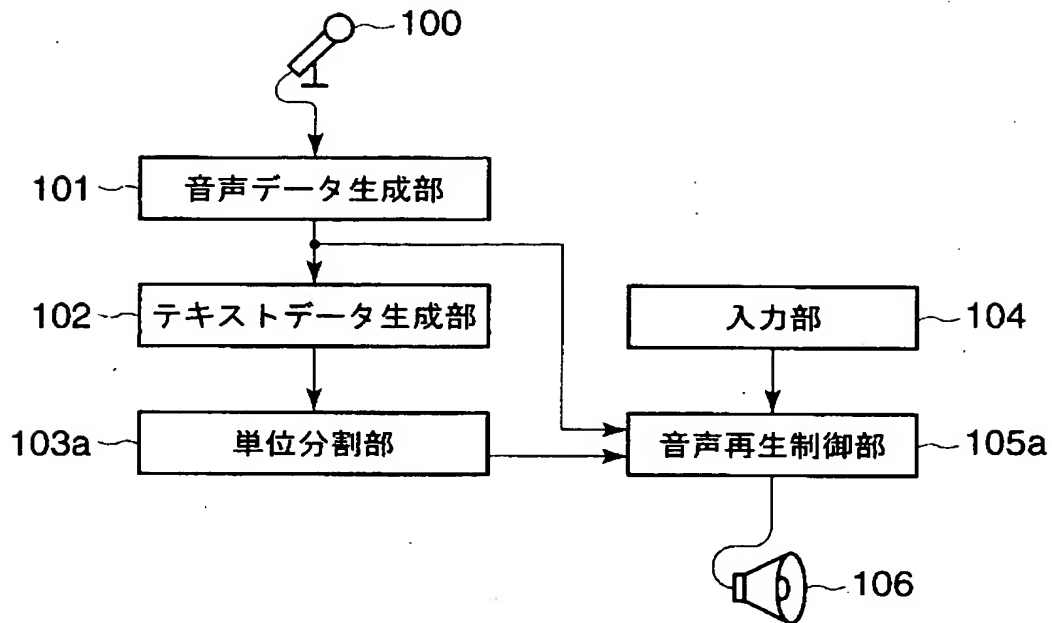
【符号の説明】

100…マイクロホン、101…音声データ生成部、102…テキストデータ生成部、103a…単位分割部、103b…単位分割部、103c…単位分割部、103d…単位分割部、103e…単位分割部、103f…単位分割部、103g…単位分割部、103h…単位分割部、103i…単位分割部、103j…単位分割部、104…入力部、105a…音声再生制御部、105b…音声再生制御部、105c…音声再生制御部、105d…音声再生制御部、105e…音声再生制御部、105f…音声再生制御部、105g…音声再生制御部、105h…音声再生制御部、105i…音声再生制御部、105j…音声再生制御部、106…スピーカ、110、120…記憶部。

【書類名】

図面

【図 1】



【図 2】

まあその一音声認識を利用することで  
△え一倍速をかなり上げても  
△十分聞き取りやすく再生できます△ (△:無音語)

【図 3】

4	6	1	4	2	2	1
ま	あ	そ	の	一	/	音
声	認	識	/	を	/	利
用	/	す	る	/	こ	と
/	で	/				
1	2	4	1	3	4	1
△	/	え	ー	/	倍	速
/	を	/	か	な	り	/
上	げ	て	/	も	/	
1	4	5	3	5	2	2
△	/	十	分	/	聞	き
取	り	/	や	す	く	/
再	生	/	で	き	/	ま
す	/	△	/			

(/:形態素区切り)

【図 4】

名詞、動詞等の自立語 : 1.0  
 形容詞、副詞等の修飾語 : 0.6  
 助詞、助動詞、活用語尾 : 0.3  
 間投詞、無音語 : 0.1

【図 5】

0.1	1.0	0.1	1.0	0.3	1.0	0.1
ま	あ	そ	の	一	/	音
声	認	識	/	を	/	利
用	/	す	る	/	こ	と
/	で	/				
0.1	0.1	1.0	0.1	0.6	1.0	0.1
△	/	え	ー	/	倍	速
/	を	/	か	な	り	/
上	げ	て	/	も	/	
0.1	0.6	1.0	0.6	1.0	0.3	0.3
△	/	十	分	/	聞	き
取	り	/	や	す	く	/
再	生	/	で	き	/	ま
す	/	△	/			



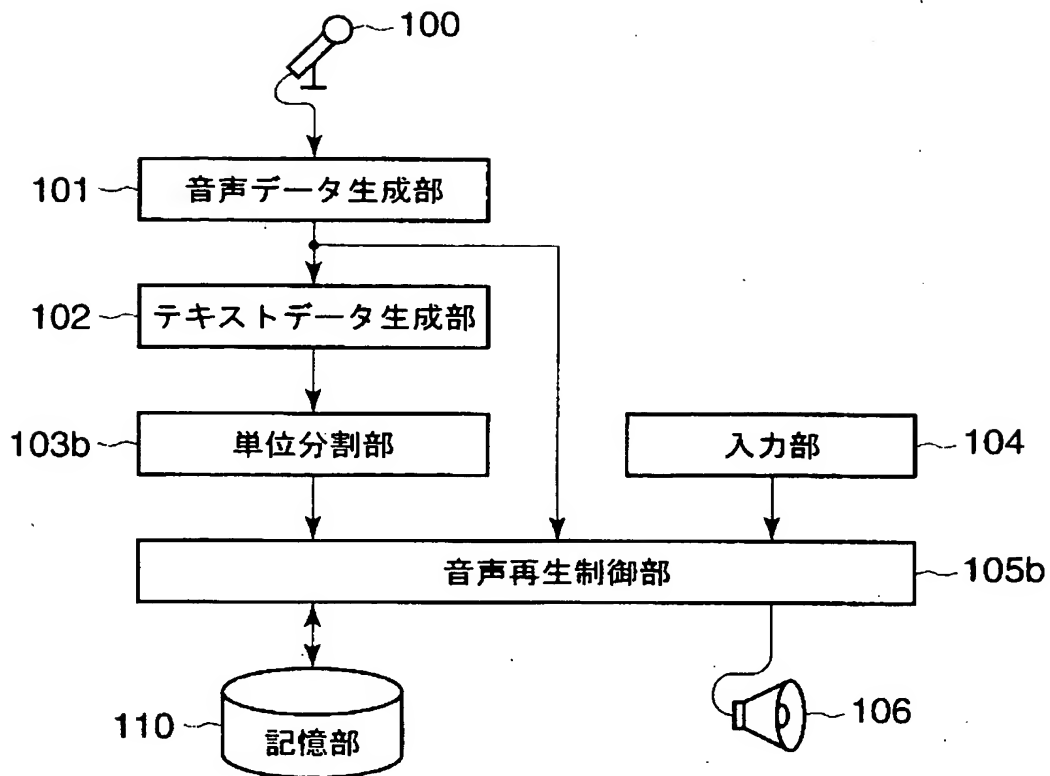
【図 6】

名詞、動詞等の自立語	: $\frac{10}{10} \times 3 \times 40 \div 60 = 2$	倍速
形容詞、副詞等の修飾語	: $\frac{10}{6} \times 3 \times 40 \div 60 = 3.3$	倍速
助詞、助動詞、活用語尾	: $\frac{10}{3} \times 3 \times 40 \div 60 = 6.7$	倍速
間投詞、無音語	: $\frac{10}{1} \times 3 \times 40 \div 60 = 20$	倍速

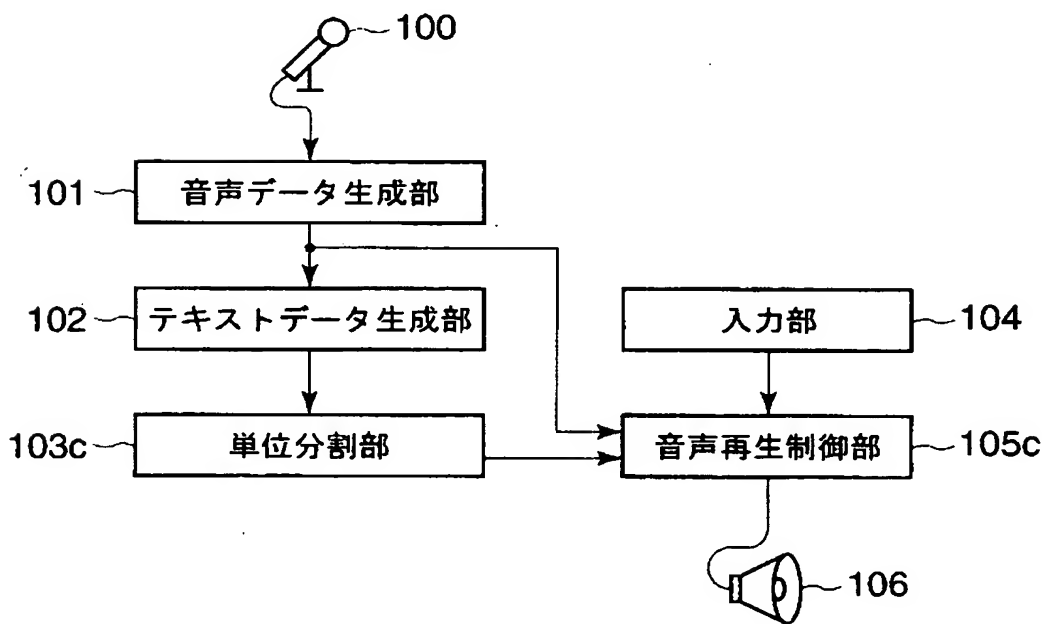
【図 7】

0.2	3	0.05	2	0.3	1	0.05
ま	あ	そ	の	一	/	音声認識 / を / 利用 / する / こと / で /
0.05	0.1	2	0.05	0.9	2	0.05
△	/	え	ー	/	倍速 / を / かなり / 上げて / も /	
0.05	1.2	2.5	0.9	2.5	0.3	0.3 0.05
△	/	十分 / 聞き取り / やすく / 再生 / でき / ます / △ /				

【図 8】



【図 9】



【図 10】

波形情報から全体の再生時間を計算して縮める方法は  
音声再生速度を高倍速にすると人間の耳では聞き取れなくなる

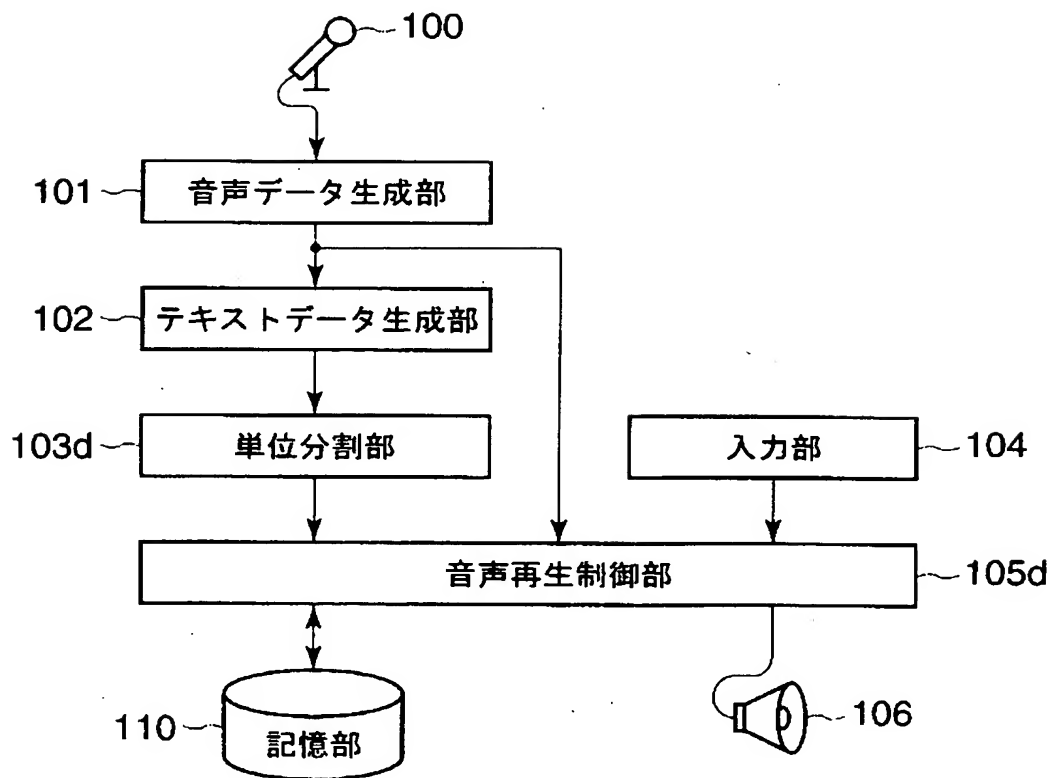
【図 1 1】

波形情報から、全体の再生時間を計算して、縮める方法は、  
音声再生速度を、高倍速にすると、人間の耳では、聞き取れなくなる、  
(：文節区切り)

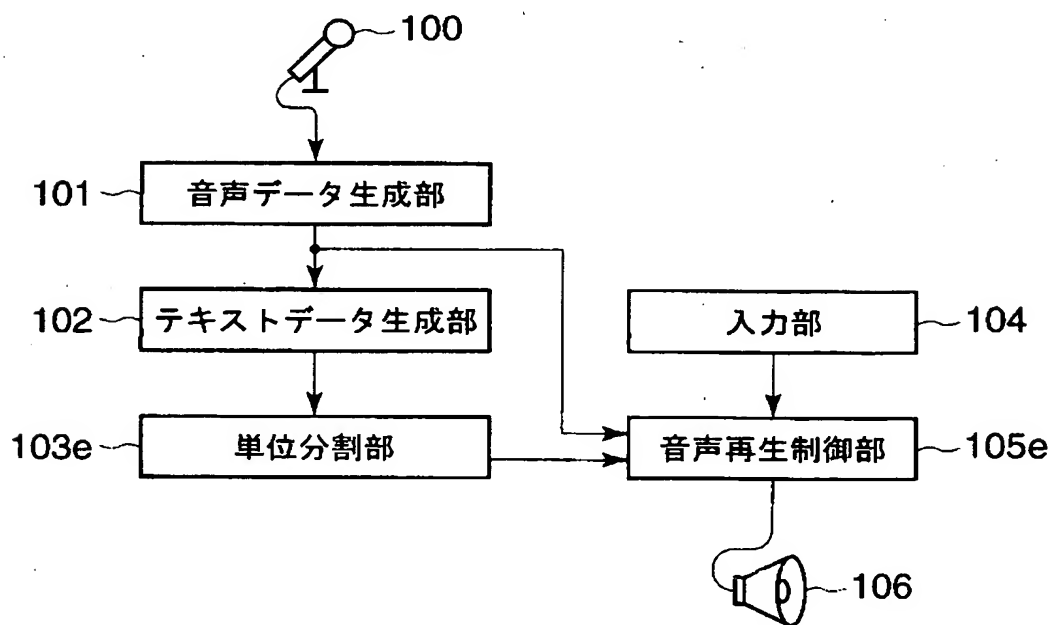
【図 12】

波形情報から、再生時間を縮める方法は、高倍速にすると聞き取れなくなる。

【図 1 3】



【図 1 4】



## 【図 15】

近年、音声単独あるいは映像と関連付けられた音声を取り扱う機器の主流は従来の磁気テープから光磁気ディスクや半導体メモリなどに移った。

その記録方式もアナログ形式からデジタル方式が広く使われるようになった。

音声の記録方式がデジタル化されることで再生時の速度の制御も容易になった。

最近のオーディオ機器やパソコン上の音声再生アプリケーションの一部にはデジタル音声の特徴を利用した再生速度可変機能が搭載されている。

これらの機器に搭載されている音声の再生速度可変機能は音声の波形データの特徴に基づいて制御する方式が使用されている。

ここでは、デジタル化された音声の波形データそのものについて一律に時間軸を短縮したり、分析したパラメータ上で全体の時間軸を短縮して再度合成したり、ある時間単位で飛び飛びに再生したりする方法が用いられている。

さらには無音部分を検出して無音部分を再生対象から削除することで全体の時間を短縮する方法も用いられる。

しかし、波形や分析パラメータ上で全体の時間軸を縮める方法では縮めすぎると人間が聞き取れなくなるためにあまり高倍速にすることはできない。

飛び飛びに再生する方法では人の喋る言葉の場合に意味が聞き取れなくなることがある。

無音部分を削除する方法は音声中の無音部分が少ない場合に効果が得られない場合がある。

これらの問題は、従来の音声再生制御方式が音声の音響的な特徴にのみ基づいた制御を行っていたことが原因である。

## 【図 16】

音声の記録方式がデジタル化されることで再生時の速度の制御も容易になった。

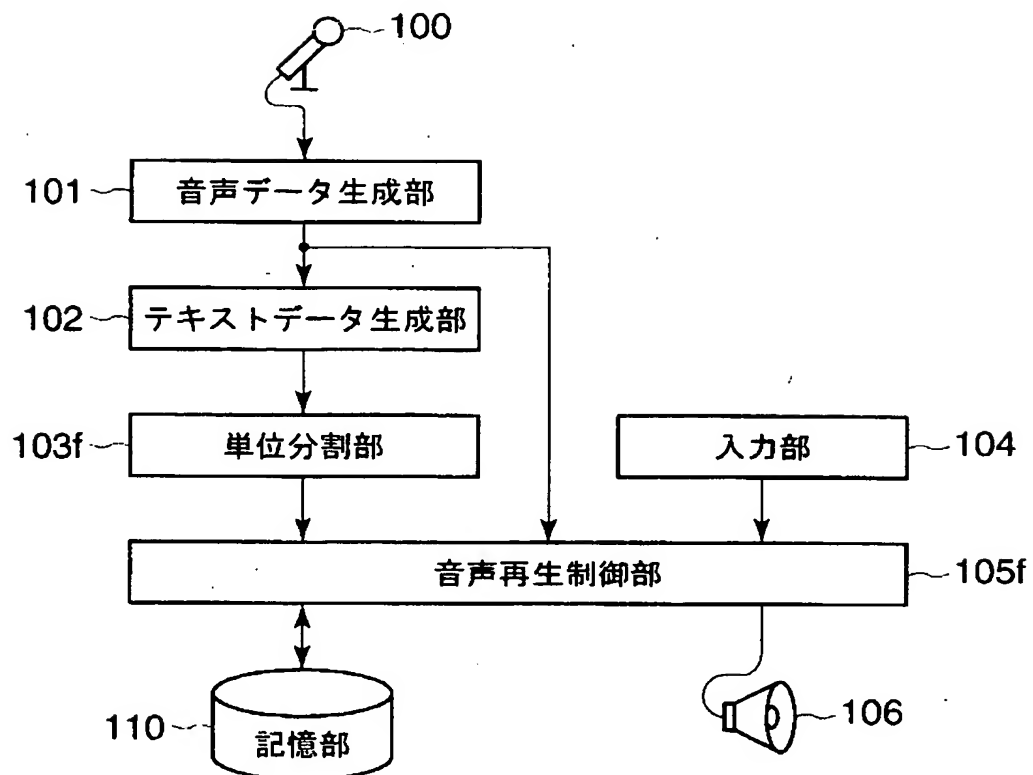
これらの機器に搭載されている音声の再生速度可変機能は音声の波形データの特徴に基づいて制御する方式が使用されている。

ここでは、デジタル化された音声の波形データそのものについて一律に時間軸を短縮したり、分析したパラメータ上で全体の時間軸を短縮して再度合成したり、ある時間単位で飛び飛びに再生したりする方法が用いられている。

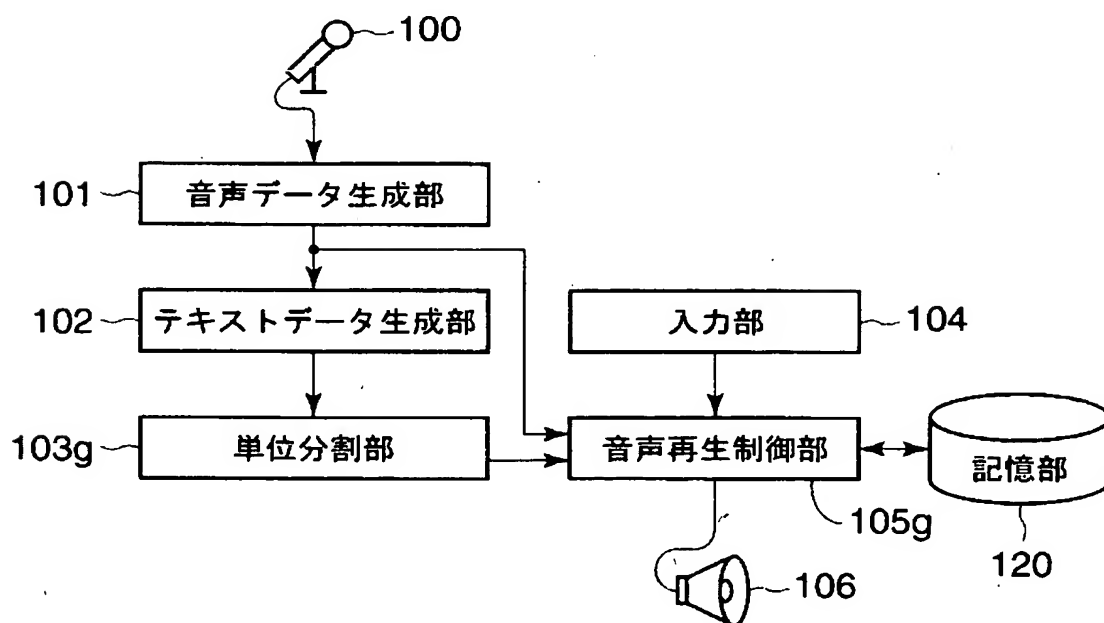
さらには無音部分を検出して無音部分を再生対象から削除することで全体の時間を短縮する方法も用いられる。

これらの問題は、従来の音声再生制御方式が音声の音響的な特徴にのみ基づいた制御を行っていたことが原因である。

【図 17】



【図 18】



【図 19】

(高)	しかし、	よって、	まず、
(中)	そして、	一方、	たたし、
(低)	例えば、	ところで、	なお、 仮に、

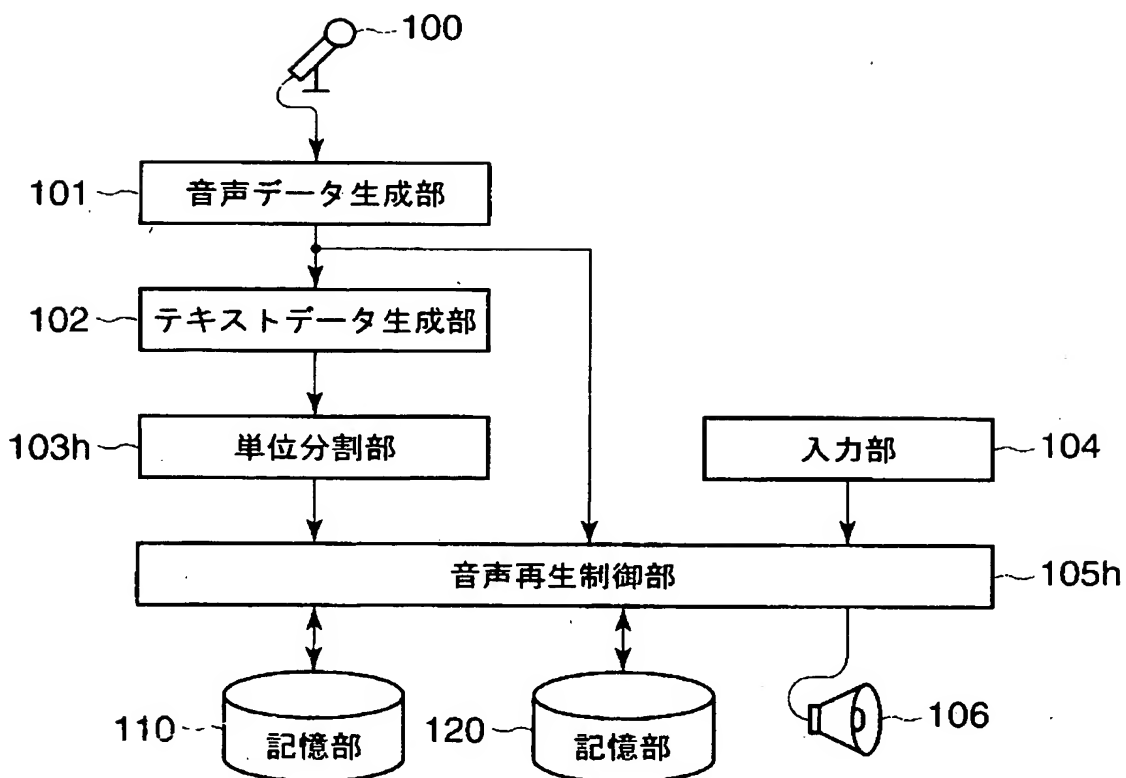
【図 20】

(1)	まず、 . . .	: 発声時間10秒
(2)	ここで、 . . .	: 発声時間10秒
(3)	そして、 . . .	: 発声時間15秒
(4)	仮に、 . . .	: 発声時間10秒
(5)	例えば、 . . .	: 発声時間10秒
(6)	しかし、 . . .	: 発声時間10秒
(7)	例えば、 . . .	: 発声時間10秒
(8)	よって、 . . .	: 発声時間 5秒
(総発声時間80秒)		

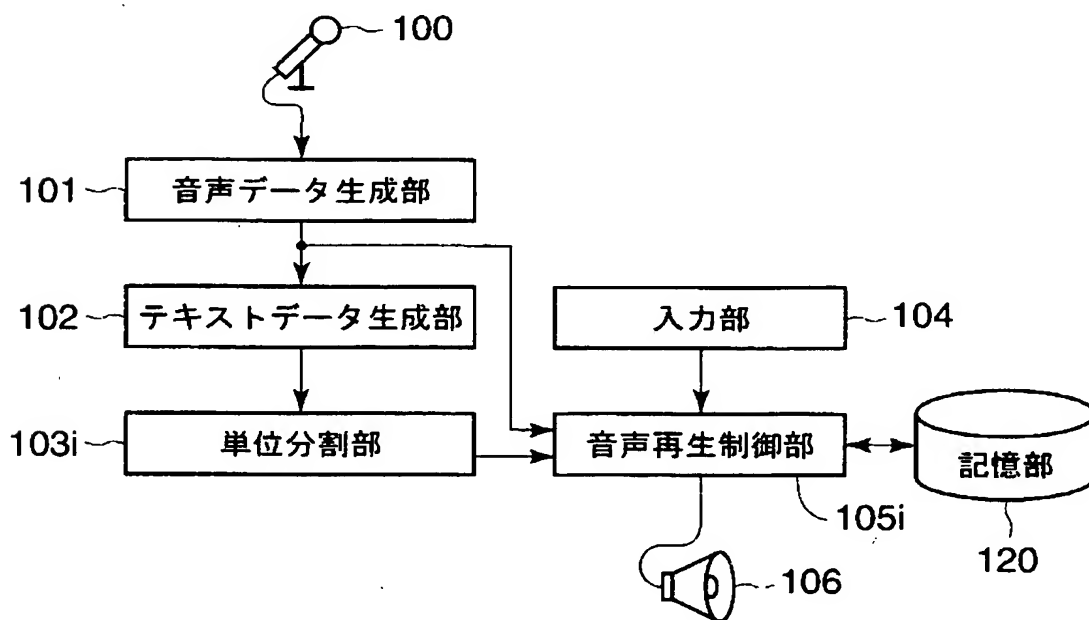
【図 21】

(1)	まず、 . . .	: 発声時間10秒
(3)	そして、 . . .	: 発声時間15秒
(6)	しかし、 . . .	: 発声時間10秒
(8)	よって、 . . .	: 発声時間 5秒
(総発声時間40秒)		

【図 2 2】



【図 2 3】





【図 2 4】

最近のオーディオ機器やパソコン上の音声再生アプリケーションの一部にはデジタル音声の特徴を利用した再生速度可変機能が搭載されている。

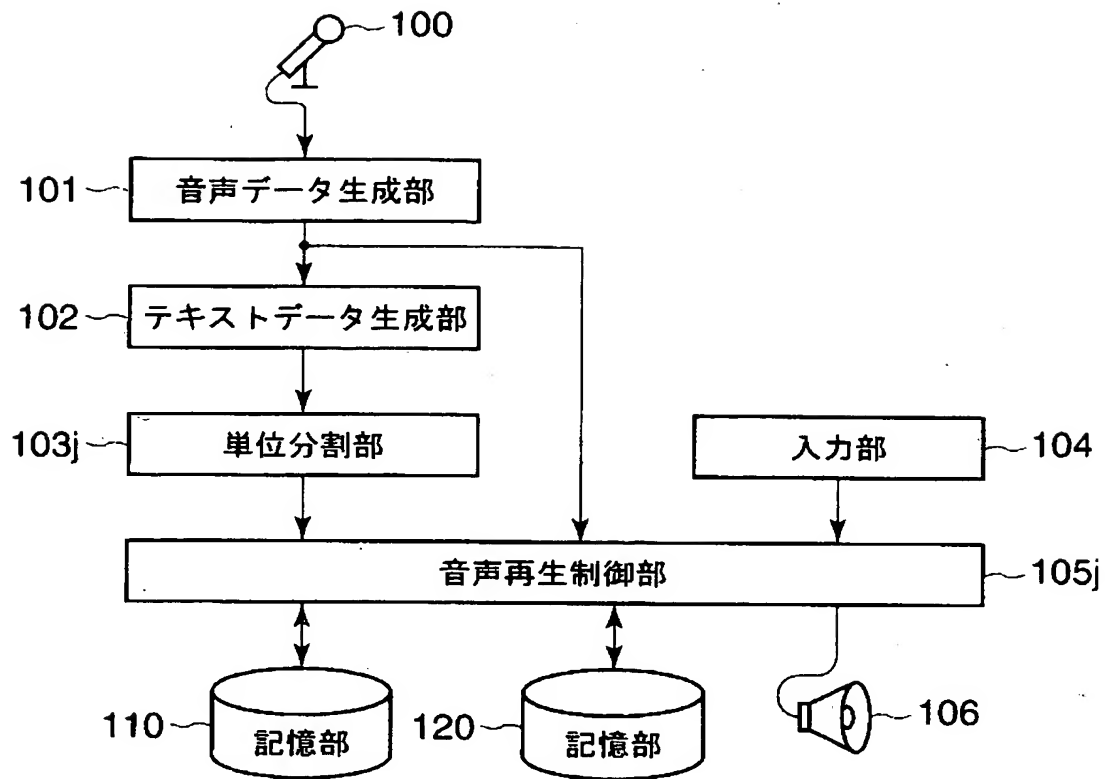
【図 2 5】

番号	単語組	確率値	発声時間(秒)
(1)	最近-の	0.3	1
(2)	の-オーディオ	0.01	3
(3)	オーディオ-機器	0.05	2
(4)	機器-や	0.1	1
(5)	や-パソコン	0.05	3
(6)	パソコン-上	0.02	2
(7)	上-の	0.1	1
(8)	の-音声再生	0.01	4
(9)	音声再生-アプリケーション	0.001	4
(10)	アプリケーション-の	0.1	1
(11)	の-一部	0.03	2
(12)	一部-には	0.05	2
(13)	には-デジタル音声	0.005	4
(14)	デジタル音声-の	0.1	1
(15)	の-特徴	0.02	2
(16)	特徴-を	0.1	1
(17)	を-利用	0.03	2
(18)	利用-した	0.1	2
(19)	した-再生速度	0.01	4
(20)	再生速度-可変機能	0.0001	4
(21)	可変機能-が	0.1	1
(22)	が-搭載	0.01	2
(23)	搭載-されている	0.1	3
	全体		52

【図 2 6】

オーディオ    音声再生アプリケーション    デジタル音声    再生速度可変機能    搭載

【図 27】



【書類名】 要約書

【要約】

【課題】 再生速度を可変したり、再生時間を短縮しても、再生される音声が表示内容を理解しやすい話速可変装置及び話速変換方法を提供する。

【解決手段】 音声データ生成部101は、マイクロホン100からの音声信号を音声データに変換する。テキストデータ生成部102は、上記音声データの発声内容を解析してテキストデータに変換する。単位分割部103aは、上記テキストデータを、言語的に意味を持つ単位に分割する。入力部104は、音声再生の速度に関わる情報として、再生倍率Nと総再生時間Lの入力を受け付ける。音声再生制御部105aは、単位分割部103aから入力されたデータと、入力部104から入力されたデータとに基づいて、分割した単位毎の再生時間を算出し、これに基づいて音声データ生成部101から入力される音声データを単位毎に音声信号に変換し、スピーカ106より拡声出力するようにしたものである。

【選択図】 図1

特願 2002-382385

出 願 人 履 歷 情 報

識別番号

[000003078]

1. 変更年月日            2001年    7月    2日  
    [変更理由]            住所変更  
                            住 所        東京都港区芝浦一丁目1番1号  
                            氏 名        株式会社東芝
  
2. 変更年月日            2003年    5月    9日  
    [変更理由]            名称変更  
                            住所変更  
                            住 所        東京都港区芝浦一丁目1番1号  
                            氏 名        株式会社東芝